

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Um sistema de apoio à decisão para um Serviço de Documentação e Informação

Pedro Manuel Lemos Graça

Mestrado Integrado em Engenharia Informática e Computação

Orientador: Gabriel Torcato David (Professor Doutor)

Julho de 2013

Um sistema de apoio à decisão para um Serviço de Documentação e Informação

Pedro Graça

Mestrado Integrado em Engenharia Informática e Computação

Aprovado em provas públicas pelo Júri:

Presidente: Prof.^a Dr.^a Ana Paula Cunha da Rocha (Professor Auxiliar)

Vogal Externo: Prof. Dr. Jorge Fernandes Rodrigues Bernardino (Professor Coordenador)

Orientador: Prof. Dr. Gabriel de Sousa Torcato David (Professor Associado)

11 de Julho de 2013

Resumo

O estudo desenvolvido no âmbito da presente dissertação surge da necessidade evidenciada pelo Serviço de Documentação e Informação (SDI) da Faculdade de Engenharia da Universidade do Porto (FEUP) em utilizar uma vasta quantidade de dados, resultante da sua atividade diária, com o objetivo de avaliar o seu desempenho e em função disso melhorar os seus processos de decisão.

Para a concretização dos objetivos propostos, foi elaborado um estudo sobre a sua estrutura orgânica, que incluiu uma análise aos processos, casos de uso e fontes de dados envolvidos nas suas atividades, culminando na construção de uma matriz que resume a arquitetura de processos do serviço, na perspetiva das fontes de dados que os suportam.

Foi ainda elaborada uma “matriz de barramento dimensional” que está na base da definição e proposta de um modelo de dados dimensional para um serviço de documentação e informação de uma instituição de ensino superior.

Posteriormente foi implementado um protótipo de um armazém de dados, baseado no modelo de dados dimensional determinado anteriormente, servindo este para avaliar a utilidade e aplicabilidade da solução proposta.

Verificou-se que o protótipo desenvolvido foi ao encontro das expectativas do serviço, verificando-se uma satisfação generalizada pelos resultados obtidos, confirmando assim a pertinência do trabalho desenvolvido e a utilidade da sua concretização.

Abstract

The purpose of this thesis arises from the need by the Documentation and Information Services of the Engineering Faculty of the Oporto University, in using a vast amount of data resulting from their daily activities, in order to assess their performance and thus improve their decision making processes.

To achieve this goal, a study to the service organic structure was carried out, which included an analysis of the processes, use cases and data sources involved in their activities, ending with the construction of a matrix that summarizes the Service's processes architecture, from the perspective of the data sources that support them.

A “dimensional bus matrix” has also been developed which is the basis of the proposed definition for a dimensional data model of an Information and Documentation Service in a Higher Education Institution.

A prototype of a data warehouse was also implemented, based on the dimensional data model given earlier, with the purpose of evaluating the utility and applicability of that solution.

The developed prototype meets the Service expectations, after an overall satisfaction being verified with the achieved results, confirming like this the relevance of the work and the usefulness of its implementation.

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao meu orientador, Professor Gabriel David. A sua sabedoria, experiência, acompanhamento e disponibilidade foram determinantes no sucesso desta dissertação.

Agradeço à Dr.^a Ana Azevedo, Diretora do Serviço de Documentação e Informação da FEUP, pela oportunidade concedida, as palavras de incentivo e o apoio que sempre manifestou.

Um enorme obrigado aos coordenadores das equipas do SDI, pelo seu inestimável contributo. Ana, Luís, Susana, Jorge, Adelaide e Cristina, sem vocês nada disto seria possível!

Aos meus colegas de trabalho um forte agradecimento, pelo constante apoio.

Um agradecimento muito especial à Palmira Seixas, à Rosário Edmonds, ao Luís Miguel e ao João Aires pelo seu apoio e auxílio nos momentos críticos.

Por último, gostaria de deixar o maior agradecimento à minha família. Nazaré, Rita, Patrícia e Tiago, Pai Manuel e Anabela, mana Sílvia, mano Vasco e Marta... por me aturarem nos melhores e piores momentos e pelo amor, companheirismo e apoio incessante. Muito obrigado!

Dedico esta dissertação a quem me é muito caro mas já não está entre nós...à minha mãe.

Sumário

Resumo.....	v
Abstract.....	vii
Agradecimentos.....	ix
Sumário.....	xi
Lista de Figuras.....	xv
Lista de Tabelas.....	xix
Abreviaturas e Símbolos.....	xxi
1 Introdução	1
1.1 Contextualização	1
1.2 Motivação e Objetivos	1
1.3 Estrutura da dissertação.....	3
2 Revisão Bibliográfica.....	5
2.1 Investigações Relevantes	5
2.1.1 <i>Bibliomining</i>	6
2.1.2 Digital Reference Electronic Warehouse (DREW) <i>project</i>	10
2.1.3 Penn Library Data Farm.....	10
2.1.4 Library Management Information System (LiMIS).....	11
2.1.5 <i>Model Warehouse</i>	12
2.1.6 Inquidoc	13
2.2 Sistemas Comerciais	15
2.2.1 Aleph Reporting Center	15
2.2.2 360.....	17
2.2.3 Journal Use Reports	19
2.2.4 ScholarlyStats.....	20
2.2.5 Ulrich's Serials Analysis System	21
2.3 Tecnologia.....	22
2.3.1 Infraestruturas de suporte à implementação de armazéns de dados	22
2.3.2 Ferramentas de suporte à modelação de armazéns de dados.....	27
2.4 Indicadores de medida.....	29
2.4.1 Indicadores para Bibliotecas	29

2.4.2	Relatórios <i>Counter</i>	30
2.5	Conclusões	30
3	Metodologia	33
3.1	Metodologia de investigação e implementação	33
4	Fase de Levantamento de Requisitos	35
4.1	Análise Orgânico Funcional do SDI	35
4.2	Processos e Casos de Uso	38
4.3	Fontes de dados	40
4.4	Matriz de Processos e Fontes de Dados	41
4.5	Conclusões	41
5	Projeto.....	43
5.1	Levantamento de indicadores.....	43
5.2	Importância e prioridade de implementação	48
5.3	Modelo dimensional.....	49
5.4	Conclusões	50
6	Implementação	53
6.1	Tecnologia.....	53
6.1.1	Justificação da adoção das tecnologias usadas.....	54
6.2	Modelo físico	54
6.3	Processo de implementação do armazém de dados.....	57
6.3.1	Criação de uma dimensão	57
6.3.2	Criação de um cubo.....	58
6.4	Extração, transformação e carregamento	59
6.4.1	Extração	60
6.4.2	Transformação.....	61
6.4.3	Carregamento	62
6.5	Particularidades de implementação.....	63
6.5.1	Implementação de tabelas de factos com dimensões degeneradas.....	63
6.5.2	Implementação de tabelas de ponte.....	64
6.5.3	Tabelas de factos que usam <i>default dimension keys</i>	67
6.5.4	<i>Slowly Changing Dimensions</i>	67
6.6	Conclusões	68
7	Apresentação de Resultados	71
7.1	<i>Oracle Business Intelligence Discoverer</i>	71
7.1.1	Arquitetura do <i>Oracle Discoverer (2009)</i>	71
7.1.2	<i>End User Layer</i>	72
7.1.3	Terminologia.....	73
7.1.4	Assistente de criação de folhas de trabalho.....	74
7.2	Análise e exploração de dados	78
7.2.1	Organização e acesso a dados	78
7.2.2	Apresentação de dados	78
7.3	Reunião de apresentação e validação de resultados	82
7.4	Conclusões	82
8	Conclusão.....	83
	Referências bibliográficas	87
	Anexos	91
Anexo A.	Bases de Dados	91
A1.	Especificações técnicas <i>MySQL</i>	91
A2.	Quadro comparativo das características de Bases de Dados	93
Anexo B.	Indicadores de medida.....	95

B.1	Norma NP ISO 11620	95
B.2	Relatórios COUNTER	96
Anexo C.	Processos e casos de uso	99
C.1	EQUIPA DE APOIO E DESCOBERTA (EAD).....	99
C.2	EQUIPA DE MEMÓRIA E ACESSO PERENE (EMAP)	102
C.3	Equipa de Representação de Informação (ERI)	104
C.4	EQUIPA DE SUPORTE À ADMINISTRAÇÃO (ESA).....	106
C.5	Equipa de Serviços Eletrónicos (ESE)	112
Anexo D.	Matriz de processos e fontes de dados	115
D.1	Matriz de processos e fontes de dados da equipa EAD.....	115
D.2	Matriz de processos e fontes de dados da equipa EMAP	115
D.3	Matriz de processos e fontes de dados da equipa ESE	116
D.4	Matriz de processos e fontes de dados da equipa AGORA	116
D.5	Matriz de processos e fontes de dados da equipa ESA	117
Anexo E.	Matriz de barramento dimensional.....	119
Anexo F.	Dicionário de factos	123
F.1	Factos dos <i>data marts</i> da equipa ERI.....	123
F.2	Factos dos <i>data marts</i> da equipa EAD	125
F.3	Factos dos <i>data marts</i> da equipa EMAP	131
F.4	Factos dos <i>data marts</i> da equipa ESE	135
F.5	Factos dos <i>data marts</i> da equipa AGORA	137
F.6	Factos dos <i>data marts</i> da equipa ESA.....	139
Anexo G.	Dicionário de dimensões	145
Anexo H.	Diagramas Entidade-Relação	155
H.1	Estrela - Documentos novos.....	155
H.2	Estrela - Documentos Alterados.....	155
H.3	Estrela – Exemplares alterados	155
H.4	Estrela – Exemplares alterados	156
H.5	Estrela - Receção de material	156
H.6	Estrela - Autoridades novas	156
H.7	Estrela – Autoridades alteradas	157
H.8	Estrela - Indexação	157
H.9	Estrela - Tratamento técnico	157
H.10	Estrela - Inscrições	158
H.11	Estrela - Circulação	158
H.12	Estrela – Requisições ao arquivo	158
H.13	Estrela - Multas	159
H.14	Estrela – Obras extraviadas	159
H.15	Estrela – Consultas ao arquivo acompanhadas	160
H.16	Estrela – Consultas ao arquivo sem acompanhamento	160
H.17	Estrela – Pedidos de apoio de 1.ª linha.....	161
H.18	Estrela – Ocupação dos pisos	161
H.19	Estrela - Fotocopiadoras.....	161
H.20	Estrela – Reserva de gabinetes	162
H.21	Estrela – Pedidos de apoio de 2.ª linha.....	162
H.22	Estrela – Participação em formação	163
H.23	Estrela - Pesquisas.....	163
Anexo I.	Código PL/SQL.....	165
I.1	Código para preencher uma tabela de ponte	165
I.2	Código alternativo para preenchimento de uma tabela de factos	167
I.3	Código para retificar números de alunos.....	169
I.4	Código para normalizar visualização de autores	171
I.5	Código para associar dimensão “Documento” à dimensão “Classe”	173

Índice Remissivo.....	175
------------------------------	------------

Lista de Figuras

Figura 2.1.1: Matriz de tópicos, perspectivas e critérios de avaliação (Nicholson 2004, 502).....	6
Figura 2.1.2: Metodologias apropriadas para cada quadrante da matriz (Nicholson 2004, 502)..	7
Figura 2.1.3: Ciclo de medições holísticas e avaliações cumulativas para Bibliotecas (Nicholson 2004)	7
Figura 2.1.4: Modelo de dados para <i>Bibliomining</i> (Nicholson 2006, 790)	8
Figura 2.1.5: Modelo conceptual para tipos de dados de um armazém de dados <i>bibliomining</i> (Nicholson 2006)	9
Figura 2.1.6: Diagrama de alto nível do Penn Library Data Farm (Zucca 2010, 6).....	11
Figura 2.1.7: Estrutura do LiMIS, baseado num armazém de dados (Decker e Höppner 2006)	12
Figura 2.1.8: Estrutura do Inquidoc.	14
Figura 2.1.9: Menu de indicadores do Inquidoc.....	14
Figura 2.1.10: Exemplo de um relatório gerado pelo Inquidoc.....	15
Figura 2.2.1: Visão conceptual do ARC. (ExLibris Group 2012, 5).....	16
Figura 2.2.2: Processo ETL e camadas ARC.....	17
Figura 2.2.3: Exemplo de relatório gerado pelo ARC. (ExLibris Group 2012).....	17
Figura 2.2.4: Núcleo fundamental para a utilização da solução 360 da Serials Solutions.	18
Figura 2.2.5: Esforço das tarefas de gestão, com e sem o 360 Counter, segundo a perspectiva do vendedor.....	18
Figura 2.2.6: Arquitetura do produto 360 da Serial Solutions. (Serial Solutions 2012a).....	19
Figura 2.2.7: O problema e respetiva solução. A visão da Swets. (Swets 2008)	20
Figura 2.2.8: Relatórios <i>ScholarlyStats</i>	21
Figura 2.3.1: Arquitetura de um armazém de dados da Microsoft (Microsoft Corporation 2012a)	24
Figura 2.3.2: Desenho de uma transformação em <i>Kettle</i> (Pentaho 2008).....	28
Figura 2.3.3: Query Studio do Cognos BI, integrado com o armazém de dados do ARC. (ExLibris Group 2012)	29
Figura 3.1.1: Metodologia de implementação do armazém de dados	34
Figura 4.1.1: Representação da estrutura orgânica formal do SDI sobreposta pela organização de trabalho transversal emanada da reestruturação de 2009.	37
Figura 4.2.1 Casos de Uso do processo Coordenação.....	39
Figura 4.2.2. Casos de Uso do processo Interpretação de Documento	39
Figura 4.2.3 Casos de Uso do processo Evento	39
Figura 6.2.1: Exemplo de implementação de uma dimensão com hierarquias	55
Figura 6.2.2: Implementação física da estrela “Circulação”	56
Figura 6.3.1: Diagrama de fluxo do processo de criação de estrelas	57
Figura 6.3.2: Assistente de configuração de dimensões. Ligação entre modelo lógico e modelo físico.....	58
Figura 6.3.3: Assistente de configuração de um cubo. Ligação entre modelo lógico e modelo físico.....	59

Figura 6.4.1: Interrogação SQL para obtenção da lista de autores de documentos.....	60
Figura 6.4.2: Exemplo de origem de dados. Lista de estudantes e seu percurso académico.	61
Figura 6.4.3: Exemplo de operação de transformação realizada numa folha <i>Excel</i>	61
Figura 6.4.4: Exemplo de <i>script</i> utilizado nas operações de transformação.	62
Figura 6.4.5: Exemplo de um mapeamento para carregamento de dados.....	62
Figura 6.4.6: Vista sobre pacote PL/SQL gerado automaticamente para mapeamento	63
Figura 6.5.1: Configuração da tabela de um cubo para suportar dimensões degeneradas	64
Figura 6.5.2: Exemplo de implementação de tabelas-de-ponte	65
Figura 6.5.3: Tipo de dados definido pelo utilizador, usado na função de agregação	65
Figura 6.5.4: Função <i>STRAGG</i>	65
Figura 6.5.5: Extração de grupos de autores	66
Figura 6.5.6: Processo de preenchimento de uma tabela de ponte.....	66
Figura 6.5.7: Diagrama entidade-relação com a implementação da estrela “Circulação”	67
Figura 6.5.8: Configurações SCD 2 para a dimensão colaborador	68
Figura 7.1.1: Arquitetura do <i>Oracle BI Discoverer</i> (ORACLE 2009).....	72
Figura 7.1.2: <i>End User Layer</i> (Rose 1995)	72
Figura 7.1.3: Áreas de negócio criadas para a apresentação do armazém de dados do SDI.....	73
Figura 7.1.4: Organização hierarquica do <i>Discoverer</i>	74
Figura 7.1.5: Seleção de objetos a apresentar numa folha de apresentação de resultados	74
Figura 7.1.6: Seleção dos itens a incluir numa folha de apresentação de resultados	75
Figura 7.1.7: Especificação de uma condição a incluir numa folha de apresentação de resultados	75
Figura 7.1.8: Especificação da organização de itens numa folha de apresentação de resultados	76
Figura 7.1.9: Especificação de parâmetros de entrada para folhas de apresentação de resultados	76
Figura 7.1.10: Especificação de ordenações para uma tabela	76
Figura 7.1.11: Interrogação SQL resultante do processo de criação de uma folha de apresentação de resultados.....	77
Figura 7.1.12: Folha de apresentação de resultados no <i>Discoverer Plus</i>	77
Figura 7.2.1: Cadernos e folhas de trabalho desenvolvidas	78
Figura 7.2.2: Apresentação de uma folha de trabalho no <i>Discoverer Viewer</i>	79
Figura 7.2.3: Utilização de formatos condicionais para análise e exploração de dados	80
Figura 7.2.4: Utilização de indicadores de aviso para análise e exploração de dados	80
Figura 7.2.5: Alteração do tipo de gráfico durante o processo de análise e exploração de dados	81
Figura 7.2.6: Navegação de dados(<i>drill down</i>)	81
Figura 7.2.7: Resultado da navegação de dados <i>drill down</i>	82
Figura C.1: Casos de Uso do processo FrontOffice	100
Figura C.2: Caso de Uso do processo Infoliteracia.....	100
Figura C.3: Casos de Uso do processo Tarefas de Gestão	101
Figura C.4: Casos de Uso do processo Gestão Documental Física.....	102
Figura C.5: Casos de Uso do processo Tratamento Técnico de Preservação.....	103
Figura C.6: Casos de Uso do processo Tarefas de Gestão (EMAP)	103
Figura C.7: Casos de Uso do processo Gestão Documental	104
Figura C.8: Casos de Uso do processo Tarefas de Gestão	105
Figura C.9: Casos de Uso do processo Coordenação.....	108
Figura C.10: Casos de Uso do processo Secretariado Técnico	108
Figura C.11: Casos de Uso do processo Gestão de Infraestruturas.....	109
Figura C.12: Casos de Uso do processo Gestão de Pessoas.....	109
Figura C.13: Casos de Uso do processo Gestão Financeira	110
Figura C.14: Casos de Uso do processo Gestão de Aquisições	110
Figura C.15: Casos de Uso do processo FEUP Edições	111
Figura C.16: Casos de Uso do processo Produção de Documentos	112
Figura C.17: Casos de Uso do processo Projetos.....	113

Figura C.18: Casos de Uso do processo Gestão de Serviços Web.....	113
Figura C.19: Casos de Uso do processo Gestão da infraestrutura tecnológica	114

Lista de Tabelas

Tabela 1: Sumário dos processos/atividades da equipa AGORA	38
Tabela 2: Tabela de fontes de dados	40
Tabela 3: Matriz processos/fontes de dados para as atividades da equipa ERI	41
Tabela 4: Lista de indicadores, distribuídos por processo e equipa	43
Tabela 5: Lista de processos acompanhada da sua importância e prioridade de implementação	48
Tabela 6: Exemplo de especificação de uma tabela de factos.....	50
Tabela 7: Exemplo da especificação de uma dimensão	50
Tabela 8: Especificações técnicas do MySQL (MySQL 2012b)	91
Tabela 9: Norma ISO11620 – Grupo B1, indicadores de perceção do utilizador.....	95
Tabela 10: Norma ISO11620 – Grupo B2, indicadores de serviços públicos.....	95
Tabela 11: Norma ISO11620 – Grupo B3, indicadores de serviços técnicos	95
Tabela 12: Norma ISO11620 – Grupo B5, indicadores de disponibilidade e utilização de recursos humanos.....	96
Tabela 13: Relatórios COUNTER para Monografias e Obras de Referência	96
Tabela 14: Relatórios COUNTER para revistas e bases de dados	96
Tabela 15: Relatórios COUNTER para Monografias e Trabalhos de Referência	97
Tabela 16: Sumário dos processos/atividades da equipa EAD	99
Tabela 17: Sumário dos processos/atividades da equipa EMAP	102
Tabela 18: Sumário dos processos/atividades da equipa ERI	104
Tabela 19: Sumário dos processos/atividades da equipa ESA (EAD (2009); FEUP Edições (2009b, 2009a))	106
Tabela 20: Sumário dos processos/atividades da equipa ESE	112

Abreviaturas e Símbolos

ACID	Atomicidade, Consistência, Isolamento, Durabilidade
AD	Armazém de Dados
ARC	<i>Aleph Reporting Center</i>
COUNTER	<i>Counting Online Usage of Networked Electronic Resources</i>
CDU	Classificação Decimal Universal
DREW	<i>Digital Reference Electronic Warehouse</i>
DML	<i>Data Manipulation Language</i>
ERP	<i>Enterprise resource planning</i>
ETL	<i>Extraction, Transformation & Load</i>
eISSN	<i>Extended International Standard Serial Number</i>
GIS	<i>Geographic Information system</i>
ISSN	<i>International Standard Serial Number</i>
ISAD	<i>International Standard Archival Description</i>
JUR	<i>Journal Usage Reports</i>
KPI	<i>Key Performance Indicators</i>
LDAP	<i>Lightweight Directory Access Protocol</i>
ROLAP	<i>Relational Online Analytical Processing</i>
MOLAP	<i>Multidimensional Online Analytical Processing</i>
OLAP	<i>Online analytical processing</i>
OLEDDB	<i>Object Linking and Embedding Database</i>
OPAC	<i>Online Public Access Catalog</i>
OWB	<i>Oracle Warehouse Builder</i>
noSQL	<i>not only SQL</i>
RAD	<i>Rapid Application Development</i>
RDBMS	<i>Relational Data Base Management System</i>
UDM	<i>Unified Dimensional Model</i>
UML	<i>Unified Modeling Language</i>
SaaS	<i>Software as a Service</i>
SCD	<i>Slowly Changing Dimensions</i>
SDI	Serviço de Documentação e Informação
SGBD	Sistema de Gestão de Base de Dados
SQL	<i>Strutured Query Language</i>
SUSHI	<i>Standardized Usage Statistics Harvesting Initiative</i>
SSD	Sistema de Suporte à Decisão

1 Introdução

Neste primeiro capítulo, apresenta-se o enquadramento e motivação desta investigação, bem como a identificação e definição do problema abordado. São resumidas as metodologias utilizadas e termina apresentando um breve resumo de cada um dos capítulos posteriores.

1.1 Contextualização

O presente documento debruça-se sobre o trabalho realizado em torno da dissertação cujo tema incide no desenvolvimento de um sistema de apoio à decisão para um Serviço de Documentação e Informação Universitário. Esse serviço é o SDI, que é a entidade da FEUP responsável por gerir, promover, disponibilizar e garantir o acesso a recursos de informação de cariz científico-técnico, pedagógico, administrativo ou patrimonial (seja ele cultural ou tecnológico).

Um Sistema de Suporte à Decisão (SSD) vem sendo definido como o conjunto de aplicações e dados necessários para apoiar na análise e no processo de tomada de decisão de uma organização (Paynter 2008). São soluções geralmente assentes em infraestruturas tecnológicas baseadas em armazéns de dados, integradas com ferramentas que permitem a análise e exploração desses dados.

Esta dissertação segue uma abordagem de estudo de caso, com um projeto concreto para o Serviço de Documentação e Informação (SDI) da Faculdade de Engenharia da Universidade do Porto (FEUP), embora se espere que o resultado do trabalho possa ter generalidade suficiente para ser útil para outros serviços similares.

1.2 Motivação e Objetivos

Hoje em dia cada vez mais as instituições necessitam de tomar decisões informadas e acertadas como forma de planearem adequadamente as suas estratégias e assim cumprirem com sucesso a sua missão.

A atual conjuntura económica mundial, sem precedentes na história, veio acentuar esta realidade, obrigando-as a diária e permanentemente decidirem bem, justificando todas as suas opções e decisões.

Os Centros de Documentação e Informação não escapam a esta realidade, competindo adicionalmente com a “já não tão recente assim” mas “cada vez mais omnipresente” *internet*, na disponibilização de recursos de informação.

Dotados de complexas estruturas organizacionais e dispondo de um variadíssimo leque de sistemas de informação de suporte às suas atividades diárias, os Serviços de Documentação e Informação são entidades geradoras de um vasto volume de dados, resultante não só das suas atividades internas, mas também da sua interação com o meio envolvente (os seus utilizadores, os seus artefactos, os seus documentos ou outros dados). Esse grande volume de dados pode ser potenciado para de alguma forma ajudar a responder adequadamente aos desafios acima formulados.

Neste contexto, o Serviço de Documentação e Informação da Faculdade de Engenharia da Universidade do Porto está empenhado em usar uma vasta quantidade de dados gerados pelos diversos sistemas e aplicações a que recorre, com o intuito de enriquecer os métodos de avaliação e gestão deste Serviço da FEUP, suportados em ajustadas definições de eficácia e exigências transparentes de eficiência. Por este motivo, esta dissertação segue uma abordagem de estudo de caso, com um projeto concreto para o SDI da FEUP, sendo advogado como ponto de partida que o desenvolvimento de uma ferramenta de apoio à decisão, baseada num armazém de dados, permitirá dar resposta aos anseios do Serviço, resolvendo os problemas formulados e expressos em seguida.

Cruzando esta realidade com o contexto previamente especificado pode claramente definir-se o problema que motiva esta dissertação e que se resume em três pontos essenciais:

- As atividades do Serviço e seus utilizadores geram um volume elevado de dados, que não são correlacionados nem tratados convenientemente;
- Existe dificuldade em avaliar o impacto da atividade do Serviço e da atividade dos seus utilizadores;
- Verifica-se uma inexistência de meios de avaliação, apoio e suporte a decisões complexas.

Baseados na motivação e problema formulados anteriormente, são enunciados em seguida os objetivos concretos desta dissertação, divididos em objetivos gerais e específicos:

- **Objetivos gerais**
 - Obter indicadores fiáveis e de qualidade;
 - Implementar um armazém de dados com vista a construir uma ferramenta de apoio à decisão e exploração dos serviços prestados pelo SDI.
- **Objetivos específicos**
 - Identificar os principais processos geridos pelo SDI;
 - Identificar os fatores críticos, critérios e aspetos relevantes para a implementação do modelo de dados dimensional de um serviço de documentação e informação de uma instituição de ensino superior;
 - Elaborar um modelo de análise do serviço com um conjunto de indicadores e gráficos;
 - Implementar um protótipo de um armazém de dados, baseado no modelo dimensional;
 - Validar o armazém de dados implementado e verificar se constitui uma ferramenta útil para a obtenção de indicadores de medida das atividades do Serviço.

Para atingir estes objetivos, esta dissertação seguirá uma abordagem baseada num estudo de caso, apoiado em metodologias próprias de um projeto de Engenharia Informática, assente numa primeira fase num modelo em cascata para a extração de requisitos e numa metodologia de

desenvolvimento iterativo, assente num processo cíclico para o desenvolvimento do armazém de dados.

1.3 Estrutura da dissertação

Para além da introdução onde se contextualiza o problema e se especificam os objetivos a atingir, este documento conta com mais 7 capítulos.

No capítulo 2, é revisto o estado da arte e são apresentados alguns trabalhos, tecnologias e normas relacionadas.

No capítulo 3, é apresentada uma descrição detalhada das metodologias de investigação e implementação adotadas.

O capítulo 4 é dedicado ao levantamento de requisitos, iniciando-se com um estudo orgânico-funcional ao SDI da FEUP, seguindo-se um levantamento de processos, respetivos casos de uso e a identificação de potenciais fontes de dados. Conclui-se o capítulo com a apresentação da matriz de processos e fontes de dados, que pretende resumir o levantamento efetuado.

O capítulo 5 é dedicado à fase de projeto, iniciando-se com a apresentação do levantamento efetuado para a recolha dos indicadores de medida pretendidos, bem como a especificação das importâncias e prioridades de implementação. A finalizar é apresentado o modelo dimensional projetado, através do seu dicionário de factos, do seu dicionário de dimensões e da sua matriz de barramento dimensional.

O capítulo 6 versa o tema da implementação. Aí são afluídos tópicos relacionados com as tecnologias adotadas, o processo de implementação e as suas particularidades.

O capítulo 7 é dedicado à apresentação de resultados, principiando com a explanação da ferramenta de suporte e finalizando com a exemplificação das potencialidades de extração, análise e exploração ao dispor dos utilizadores do armazém de dados.

Para finalizar, o capítulo 8, apresenta as conclusões finais da presente dissertação e aponta sugestões para trabalho futuro relacionadas com a temática aqui estudada e abordada.

2 Revisão Bibliográfica

A informação apresentada neste capítulo está dividida em cinco secções:

Na primeira secção serão apresentadas algumas investigações relevantes, aplicadas às áreas dos Sistemas de Suporte à Decisão e dos Armazéns de Dados, tendo sempre como horizonte comum Serviços de Documentação e Informação Universitários.

Na segunda secção apresentam-se exemplos de soluções comerciais aplicadas aos domínios dos SSD e AD, em contexto de Serviços de Documentação e Informação Universitários.

A terceira secção é dedicada à revisão tecnológica, onde será apresentado o panorama atual em matéria de ferramentas ao dispor para a construção de SSD e AD.

Na quarta secção serão apresentadas normas sobre indicadores de medida normalmente utilizados em Serviços de Documentação e Informação e que eventualmente terão alguma aplicabilidade prática neste contexto.

Por fim, na quinta e última secção, serão expostas algumas conclusões que evidenciam as lacunas existentes, sublinhando desta forma a relevância desta investigação.

2.1 Investigações Relevantes

Esta secção incide sobre a prospeção realizada para a descoberta de trabalhos de investigação relevantes para a investigação em causa, que utilizassem tecnologias e técnicas associadas ao domínio dos Armazéns de Dados, aplicadas à área dos Serviços de Documentação e Informação Universitários.

Salienta-se que muita da literatura investigada tinha como finalidade a aplicação de técnicas de mineração de dados, não se centrando especificamente na análise, investigação e construção de armazém de dados, nem nas técnicas habitualmente associadas a esta área de investigação, como seja a análise e exploração de possíveis modelos de dados.

Pelo exposto anteriormente e pressupondo que normalmente os armazéns de dados são as estruturas que suportam os dados que serão posteriormente objeto da aplicação de técnicas de mineração, não se deixa de apresentar esses mesmos trabalhos, dando tanto quanto possível enfoque aos armazéns de dados.

2.1.1 Bibliomining

“*Bibliomining*” foi um termo criado e introduzido por Nicholson e Stanton (2003) no decurso do trabalho de investigação realizado sobre o emprego de técnicas de mineração de dados ao domínio das bibliotecas. No contexto da literatura de investigação, a maioria dos trabalhos que incluíam os termos *library* e *data mining* não se referiam a dados de bibliotecas tradicionais, mas antes a bibliotecas de *software*, dado que *data mining* é a aplicação de técnicas de uma grande biblioteca de ferramentas. A fim de tornar a pesquisa mais propícia para aqueles mais preocupados com a mineração de dados numa biblioteca tradicional, o termo *bibliomining* foi criado. O termo é uma homenagem à bibliometria, que é a ciência da descoberta de padrões na comunicação científica (Nicholson 2003).

Bibliomining é pois a aplicação de ferramentas de mineração de dados e bibliometria, através de um armazém de dados, a dados originários dos serviços oferecidos pela Biblioteca e da atividade dos seus utilizadores, com o objetivo de auxiliar a tomada de decisão e a justificação de serviços (Prakash, Chand, e Gohel 2004; Shieh 2010; Pal 2011).

Baseado no conceito de *Bibliomining*, Nicholson realizou múltiplos ensaios de investigação que inclusive inspiraram outros tantos estudos de autores diversos.

Um dos primeiros trabalhos realizados por Nicholson, após a instituição do termo, foi o estudo onde era caracterizado o chamado processo de *Bibliomining*, que compreendia as áreas de armazenamento e mineração de dados e ainda a área de apoio à decisão em Bibliotecas.

Segundo Nicholson (2003) o processo de *Bibliomining* consiste nas seguintes tarefas:

- Determinação de áreas de enfoque;
- Identificação de fontes de dados internas e externas;
- Recolha, limpeza e anonimato dos dados de um armazém de dados;
- Seleção de ferramentas de análise adequadas;
- Descoberta de padrões através de mineração de dados e criação de relatórios com as ferramentas tradicionais de análise;
- Análise e aplicação dos resultados.

A propósito deste processo Shieh (2010, 710), afirma que a construção do armazém de dados é o fator crítico que determina o sucesso da aplicação do *Bibliomining*.

Mais tarde Nicholson (2004), desenvolve trabalho focado na área do suporte à tomada de decisão, através da definição de uma base conceptual para a medição e avaliação de serviços de Biblioteca.

Nesta investigação Nicholson, faz a distinção entre medição e avaliação, afirmando que a “medição é apenas um precursor da avaliação” (Nicholson 2004, 497) e que “a medição produz dados; no entanto a avaliação produz informação” (Nicholson 2004, 502).

Desenvolve uma matriz de indicadores de medida (Figura 2.1.1), que demonstra a relação entre tópicos, perspectivas e critérios de avaliação.

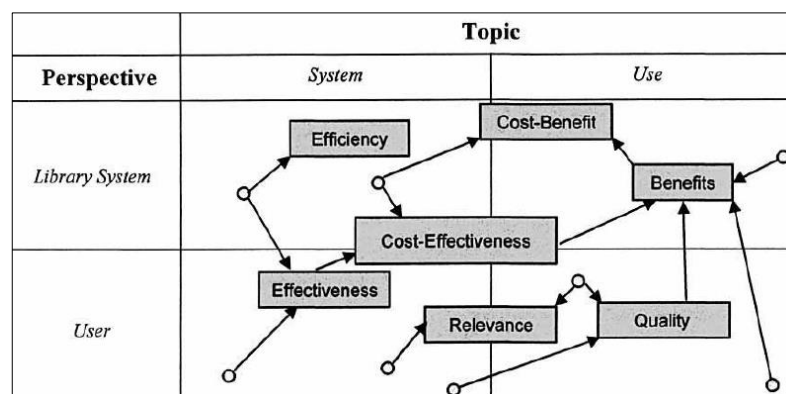


Figura 2.1.1: Matriz de tópicos, perspectivas e critérios de avaliação (Nicholson 2004, 502).

Associada à matriz, Nicholson sugere também algumas das metodologias que podem ser usadas (Figura 2.1.2). Note-se a inclusão do *bibliomining* como metodologia apropriada para o quadrante correspondente ao ponto de vista interno do uso.

Perspective	Topic	
	Library System	Use
Internal (Library System)	Staff surveys and interviews Audits of collections, systems, or staff	Bibliomining Transaction / Web Log Analysis Observation of user behavior
External (User)	Surveys and interviews Talk-alouds and in-process feedback mechanisms Focus groups	Surveys and interviews Focus Groups User citation tracking

Figura 2.1.2: Metodologias apropriadas para cada quadrante da matriz (Nicholson 2004, 502)

Segundo Nicholson (2004, 505), muitos processos de tomada de decisão nas Bibliotecas, têm origem em problemas e a sua avaliação não é feita através de uma análise sistemática das situações.

A organização esquematizada na Figura 2.1.3 pode ser usada como um instrumento sólido de resolução de problemas, em detrimento de uma solução temporária e ineficiente, proporcionando uma compreensão completa e global da Biblioteca através da avaliação de numerosos indicadores, analisados de diferentes perspectivas, garantindo uma eficiente tomada de decisão.

Ainda com o *bibliomining* como tema central, Nicholson (2006) explora a integração da extração de dados e dos serviços digitais de Bibliotecas, aprofundando primeiro o conceito, e posteriormente estabelecendo a comparação do processo de *Bibliomining*, do ponto de vista dos decisores de Bibliotecas e dos investigadores na área. Finalmente desenvolve uma agenda de investigação com o intuito de promover a resolução de muitos dos problemas frequentes no *Bibliomining*, de forma a impulsionar a investigação.

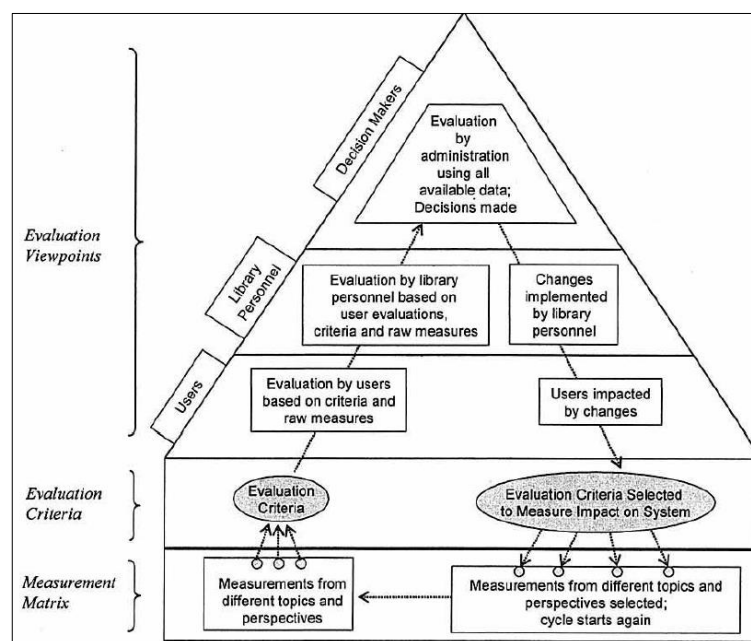


Figura 2.1.3: Ciclo de medições holísticas e avaliações cumulativas para Bibliotecas (Nicholson 2004)

Neste trabalho, Nicholson (2006) aflora o tema dos armazéns de dados para suportar *bibliomining*. Afirmar que esses armazéns de dados têm como fontes os dados bibliométricos e os

dados sobre o uso dos sistemas *web*, que por sua vez têm como campo comum – a obra acedida. Afirma ainda que “combinando estas duas fontes de dados num armazém de dados permitirá que investigadores e gestores de Bibliotecas possam entender melhor o espaço de informação que eles criaram” (Nicholson 2006, 789).

A Figura 2.1.4 mostra o modelo de dados para *bibliomining*, demonstrando o conceito de criação de ligações entre obras, baseado no processo de criação (bibliometria) e no acesso a essas obras (Nicholson 2006, 790).

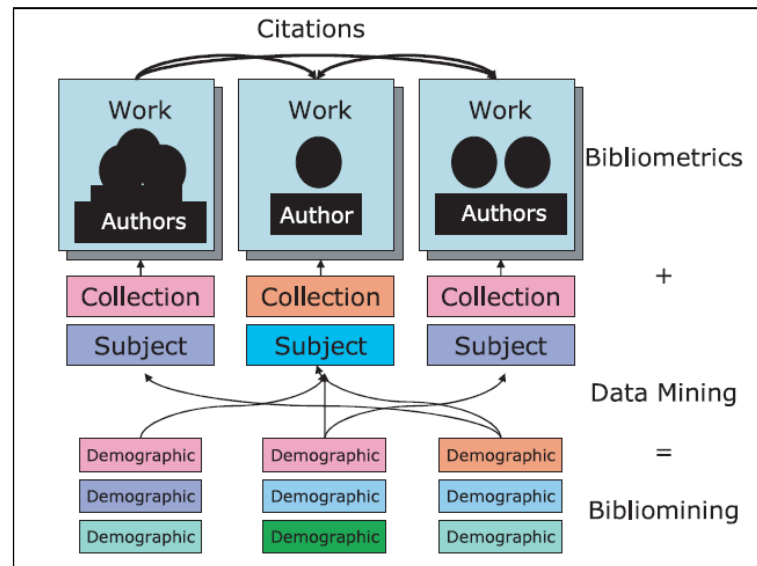


Figura 2.1.4: Modelo de dados para *Bibliomining* (Nicholson 2006, 790)

Ainda ligado ao tema dos dados do armazém, Nicholson (2006, 790-791) propõe e defende a utilidade de um modelo conceptual para a determinação dos campos de dados dos sistemas operacionais que são importantes preservar.

Esse modelo é composto por três áreas:

- Área de dados sobre a obra – este tipo de dados contém três tipos de campos: campos extraídos da obra (ex. título e autor), campos criados sobre a obra (ex. classificação) e campos que indicam o formato e localização da obra (ex. URL ou coleção). Esta área pode conter também informação bibliométrica, como citações ou ligações a outras obras. Nicholson, sobre este assunto refere as dificuldades de contabilizar acessos ao nível do artigo devido ao facto de muitos fornecedores reportarem a informação sobre o uso de recursos ao nível do título da obra. Alerta para a necessidade de serem criados indicadores semelhantes aos agregadores COUNTER¹ para o uso de recursos, ao nível do artigo;
- Área de dados sobre o utilizador – este tipo de dados contém campos relacionados com informação demográfica do utilizador. Adicionalmente outros campos provenientes da atividade do utilizador que possam ser inferidos são armazenados aqui (ex. o endereço IP);
- Área de dados sobre o serviço – este tipo de dados é o mais difícil de conceptualizar, dado haver muitos tipos de serviços (pesquisa, circulação, referência, empréstimo interbibliotecário), todos eles com campos comuns e específicos que podem ser capturados no armazém de dados.

A Figura 2.1.5 demonstra o modelo conceptual para os dados do armazém.

¹ COUNTER: Indicadores que medem a utilização recursos eletrónicos (<http://www.projectcounter.org/>)

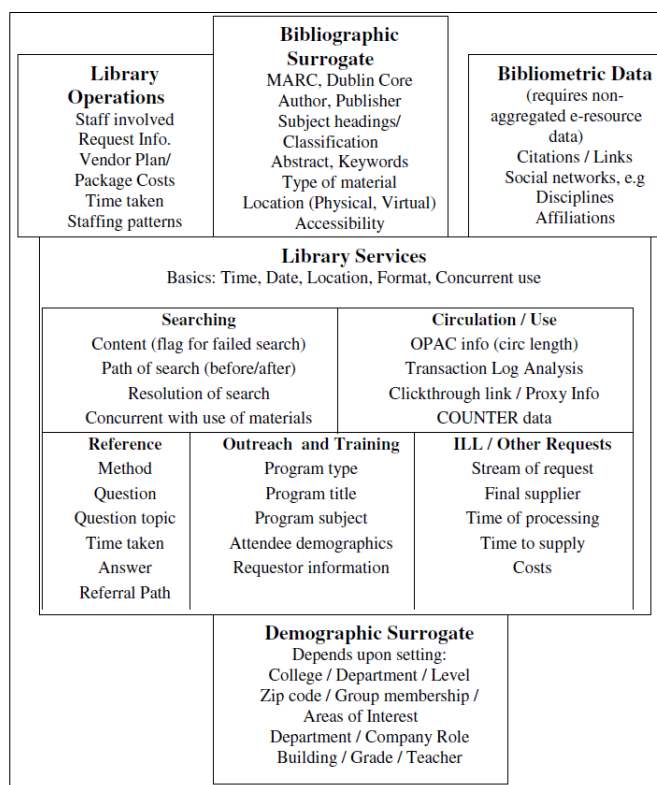


Figura 2.1.5: Modelo conceptual para tipos de dados de um armazém de dados *bibliomining* (Nicholson 2006)

No seu trabalho de investigação Nicholson (2006, 791-795), elenca ainda a utilidade de um armazém de dados baseado no conceito *bibliomining*, fazendo paralelamente menção a um projeto desenvolvido por Joe Zucca na Universidade da Pensilvânia (o *Penn Library Data Farm Project*) e ainda ao seu próprio projeto, o *Digital Reference Electronic Warehouse (DREW) project*, em colaboração com David Lankes (Nicholson e Lankes 2007). Estes dois projetos serão ainda objeto de análise no decorrer desta investigação.

Quanto à utilidade de um armazém de dados baseado no conceito *bibliomining*, esta é elencada através dos seguintes pontos:

- Utilização de relatórios tradicionais, onde listas, agregações, médias e outras medidas podem ser geradas a partir dos dados recolhidos ao longo dos tempos;
- Utilização de ferramentas OLAP (*OnLine Analytical Processing*), método através do qual é possível analisar dados com a ajuda de um interface que facilita a navegação por diferentes variáveis, tipos de agregação, linhas temporais ou dimensões;
- Utilização de métodos e ferramentas de visualização para explorar dados, permitindo a rápida descoberta de padrões;
- Emprego de técnicas de extração de dados, incluindo técnicas referidas anteriormente e também outras associadas às áreas da estatística e inteligência artificial, com o objetivo de descobrir padrões de conhecimento novos e úteis;
- Estabelecer a ponte entre investigadores e profissionais, tendo o armazém de dados como ponto de colaboração virtual.

2.1.2 Digital Reference Electronic Warehouse (DREW) project

Com este projeto Nicholson e Lankes (2007), criaram uma plataforma colaborativa que fomenta a investigação na área da criação de bases de conhecimento relacionadas com serviços de referência digital.

O projeto DREW tem como objetivo a criação de um armazém de dados para transações de serviços de referência, que potencie a investigação e que possibilite a criação de ferramentas de análise e avaliação utilizáveis pelos gestores de serviços de referência (Nicholson e Lankes 2007, 46).

Para atingir este objetivo, o projeto DREW desenvolve um modelo de dados que permite que transações digitais de diferentes serviços de referência e de formatos de comunicação diferentes coabitem no mesmo armazém de dados (Nicholson 2006, 795).

Este projeto utiliza alguns dos conceitos de *bibliomining* apresentados na secção anterior. Constitui o espaço ideal para o desenvolvimento e aprofundamento do conceito, assim como para a criação de novas ferramentas de medição e descoberta de padrões, tendo em conta a grande quantidade de dados aí disponíveis. Ainda outro objetivo é a criação de um Sistema de Gestão de Informação que possa ser aplicado a toda a base de dados, para fins de pesquisa e para que as bibliotecas participantes possam aceder para saber mais sobre os seus próprios serviços (Nicholson e Lankes 2007, 57).

Concluindo, este armazém de transações de referências digitais permitirá um nível de compreensão sobre os serviços disponibilizados pelas Bibliotecas, anteriormente indisponíveis a investigadores e educadores. Adicionalmente, os administradores de serviços participantes terão acesso a relatórios personalizados e ferramentas de gestão de informação à medida que elas vão sendo desenvolvidas (Nicholson e Lankes 2007, 57).

2.1.3 Penn Library Data Farm

O armazém de dados da Biblioteca da Universidade da Pensilvânia - *Penn Library Data Farm* - é apontado por Nicholson (2006, 793) como um projeto exemplar, sendo considerado um dos mais avançados projetos baseados no conceito de *Bibliomining*. Este projeto é um armazém de dados com uma interface que permite a possibilidade de explorar os dados de uso da Biblioteca, estando disponível a todos os utilizadores interessados. No sítio web deste projeto (University of Pennsylvania Library) é possível também depositar e partilhar ficheiros de dados relativos a estatísticas de uso de Bibliotecas

A Figura 2.1.6 exemplifica o processo de identificação, extração, transformação e carregamento de dados do *Penn Library Data Farm*. O início do procedimento começa pela identificação de potenciais fontes de dados (representado pelo círculo verde na figura), segue-se o carregamento desses dados e de outros relacionados com os utilizadores e a comunidade. No ambiente do armazém de dados (círculo a azul na figura), existem processos que capturam esses dados e os limpam, normalizam e os tornam anónimos, depositando-os no repositório central (que é uma base de dados relacional). Posteriormente são construídos vários tipos de ferramentas sobre esse repositório, que possibilitam a interação dos utilizadores com os dados. O trabalho analítico dos utilizadores e as suas decisões repercutem-se outra vez no ambiente do serviço por meio de planeamento, desenvolvimento de coleções, afetação de pessoal e outros processos.

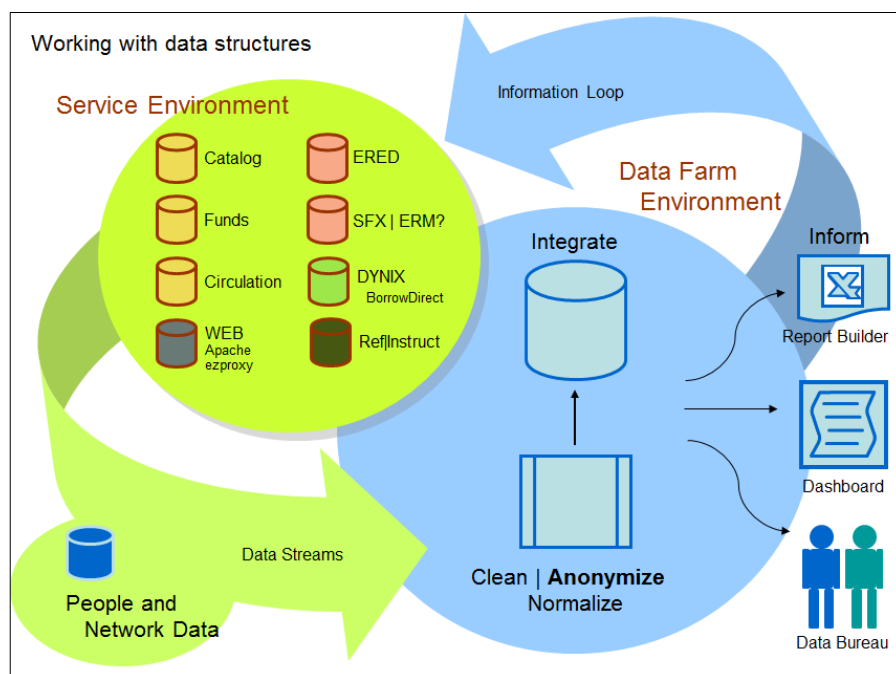


Figura 2.1.6: Diagrama de alto nível do Penn Library Data Farm (Zucca 2010, 6)

De referir que o projeto *Penn Library Data Farm* é suportado por uma ferramenta chamada *Metridoc*, em desenvolvimento na própria Universidade da Pensilvânia que permite a integração de dados e providencia suporte na construção de infraestruturas de apoio à decisão e planeamento (University of Pennsylvania Library 2012).

2.1.4 Library Management Information System (LiMIS)

O *Library Management Information System* (LiMIS) surge na sequência das discussões levadas a cabo na Biblioteca da Universidade de Bielefeld sobre eficácia e eficiência. A sua origem advém da investigação de Decker e Höppner (2006), que para além de se concentrar nos requisitos e condições necessárias para a implementação de gestão de conhecimento sobre o comportamento dos utilizadores em Bibliotecas Universitárias, também propõe um modelo conceptual para um sistema de gestão de informação de Bibliotecas, para suporte de processos de planeamento estratégico, baseado num armazém de dados.

Em termos de planeamento estratégico, a reflexão levada a cabo na Biblioteca da Universidade de Bielefeld iniciou-se com a recolha de ideias para novos serviços, contando com a colaboração fundamental dos seus colaboradores, de peritos externos e ainda de literatura relacionada com o tema. Este exercício permitiu construir 50 folhas de novos produtos/serviços, classificadas quanto à sua importância óbvia para os utilizadores e que foram confrontadas com os serviços existentes, permitindo assim estabelecer um panorama geral do estado da oferta de serviços da Biblioteca.

Ainda relacionado com o processo reflexivo sobre planeamento estratégico levado a cabo na Biblioteca de Universidade de Bielefeld, Decker e Höppner (2006, 506) concluem que somente a eficácia tinha sido considerada, faltando lidar com a eficiência. Quanto à eficiência dos serviços, os tópicos mais importantes a considerar foram: todas as atividades relacionadas com o apoio ao utilizador, incluindo o Marketing, devido à noção tradicional de bibliotecas detida pelos seus utilizadores; a avaliação crítica dos processos de trabalho e a sua relação com os indicadores de desempenho instituídos na Biblioteca que aponta para algumas deficiências já conhecidas, mas não completamente entendidas; por último, mas não menos importante, a necessidade de verificar

as rotinas de processamento de dados, que são um fundamento indispensável para a maioria dos serviços, mas que entretanto ganharam grande complexidade.

Deste modo, o processamento de dados foi também considerado (Decker e Höppner 2006, 506) como um alicerce fundamental para o planeamento estratégico, onde sofisticadas técnicas de análise de dados são necessárias para permitir o suporte à decisão no contexto de bibliotecas académicas.

A gestão de conhecimento é o produto da aplicação dessas técnicas, sendo considerado segundo Decker e Höppner (2006) uma peça fundamental no processo de suporte à decisão e planeamento estratégico, que por sua vez assume, cada vez mais, um papel crucial nas Bibliotecas, para justificar a sua existência e justificar as suas opções. Neste sentido, uma infraestrutura de suporte, gestão e criação desse conhecimento assume, por inerência, uma importância igualmente fundamental.

Não existe consenso unânime sobre o conceito de criação e gestão de conhecimento nem um conjunto bem definido de técnicas na comunidade especializada. Por outro lado, a existência de um armazém de dados, bem como um conjunto de métodos (onde se inclui o OLAP e a mineração de dados), parece oferecer um certo consenso generalizado no que a instrumentos necessários para a sua concretização diz respeito (Decker e Höppner 2006, 507).

A Figura 2.1.7 mostra a estrutura básica de um sistema de gestão de informação para Bibliotecas, no contexto da gestão e criação de conhecimento.

A peça central deste sistema é um armazém de dados e a ideia subjacente é transferir os dados gerados pelos diferentes sistemas operacionais da Biblioteca para esse armazém de dados. Durante o processo de transferência de dados são aplicadas as transformações necessárias aos dados, com o objetivo de os limpar, de os normalizar e de os anonimizar. Na figura pode também verificar-se a presença de instrumentos de análise geridos por uma interface.

Resumindo, um armazém de dados e as técnicas de OLAP/Mineração de Dados constroem a ponte entre os “dados de utilizadores” e a “informação de utilizadores”, enquanto o “conhecimento sobre utilizadores” é o resultado da tradução e disseminação de “informação de utilizadores” na biblioteca (Decker e Höppner 2006, 509).

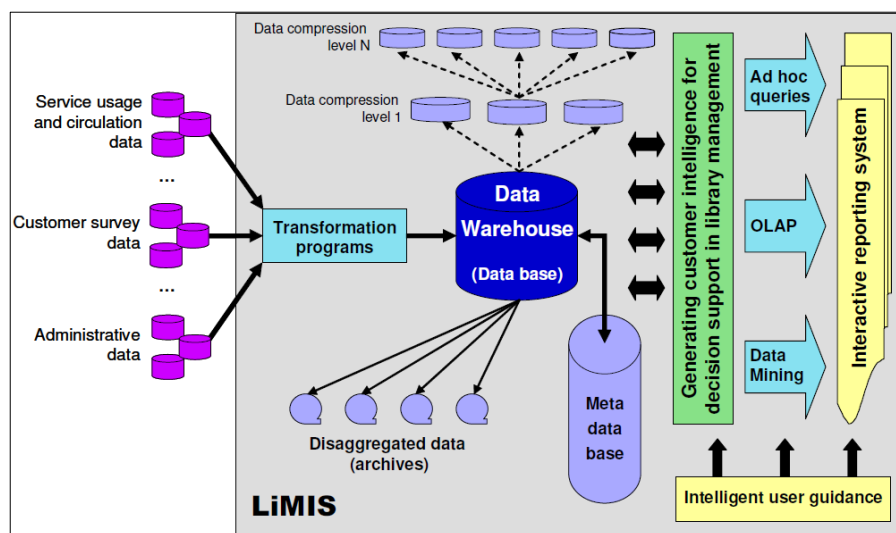


Figura 2.1.7: Estrutura do LiMIS, baseado num armazém de dados (Decker e Höppner 2006)

2.1.5 Model Warehouse

O trabalho desenvolvido por Awan e Daneshgar (2011), da Universidade de Nova Gales do Sul, é mais um exemplo do trabalho de investigação que relaciona técnicas e conceitos de armazéns de dados e a criação e gestão de conhecimento sobre clientes, aplicado ao contexto das Bibliotecas Académicas.

O estudo defende que a gestão de conhecimento sobre clientes, assistida por uma taxonomia apropriada, potenciará melhorias na atual oferta de serviços disponíveis ao utilizador, bem como a descoberta e desenvolvimento de novos e inovadores serviços nas Bibliotecas Académicas. Defende que a implementação de um sistema ajudará as Bibliotecas a manter a liderança como centros de criação e disseminação de saber e a ultrapassar os desafios que hoje em dia se lhes apresentam, como seja a competição com a *internet*, que garante acesso simplificado a uma enorme massa de informação (Awan e Daneshgar 2011, 352).

Este estudo propõe uma metodologia de alto nível para a implementação de um Armazém de Modelos, como primeiro passo para atingir o objetivo anteriormente proposto.

Um Armazém de Modelos ou Armazém de Conhecimento é similar a um armazém de dados, mas na sua essência armazena conhecimento, ideias e experiências (Awan e Daneshgar 2011).

A metodologia de alto nível proposta envolve: a seleção de um modelo teórico para o conhecimento de clientes, a partir de literatura existente; envolve a extração, categorização e armazenamento de conhecimento, no armazém de modelos desenvolvido, segundo o modelo teórico anteriormente escolhido, com o auxílio de um *software* de análise de conteúdos; envolve a análise e síntese das várias categorias de conhecimento de clientes com o objetivo de criar lógica de negócio, levando à descoberta e desenvolvimento de novos serviços ou produtos; por fim envolve a validação e verificação de resultados, assim como a avaliação dos benefícios da taxonomia de conhecimento proposta e o armazém de modelos associado (Awan e Daneshgar 2011, 354).

O modelo de dados deste armazém segue a taxonomia da base de conhecimento, sendo apelidada por Awan e Daneshgar (Awan e Daneshgar 2011, 354) como o “modelo de conhecimento” apropriado para desenvolver o armazém de dados.

2.1.6 Inquidoc

O Inquidoc é um sistema de apresentação de indicadores de desempenho para serviços de Biblioteca ou Centros de Documentação.

O desenvolvimento desta aplicação surgiu no âmbito de um estágio curricular da Licenciatura de Ciência da Informação, curso híbrido, ministrado pela Faculdade de Engenharia, em parceria com a Faculdade de Letras da Universidade do Porto.

O principal objetivo da aplicação é a transformação de dados em bruto em informação, isto é, a modelação de dados em bruto dos sistemas de informação das bibliotecas, dispersos na maior parte dos casos por vários sistemas de informação, numa aplicação central que possibilite a geração dos indicadores de desempenho para as bibliotecas e centros de documentação de acordo com a norma ISO 11620 (ver a secção 2.4.1) (Neves 2007, 15).

O Inquidoc utiliza tecnologia Microsoft no seu desenvolvimento, mais propriamente Microsoft Access, uma ferramenta RAD.

Utiliza ainda Oracle como SGBD (sistema gestão de base de dados), no servidor do sistema de gestão integrada de bibliotecas (Neves 2007, 17-18).

Para a obtenção dos dados necessários à geração dos indicadores da norma 11620, o Inquidoc utiliza as metodologias comumente empregues em armazéns de dados.

O processo de extração dos dados é efetuado diretamente da base de dados do sistema integrado de gestão de bibliotecas, utilizando ferramentas de interoperabilidade que são propriedade da Microsoft, mais precisamente a ligação à base de dados Oracle é efetuada com base em OLEDB. Os dados são manipulados com base em SQL, que extraem apenas a informação que é relevante para a retirada dos indicadores dos sistemas integrados de gestão de bibliotecas para o Inquidoc. Os processos de extração de dados foram desenhados de forma a eliminar os dados que potencialmente levassem à obtenção de resultados errados ou inconsistentes (Neves 2007, 20).

A Figura 2.1.8 apresenta uma visão geral da estrutura do Inquidoc.

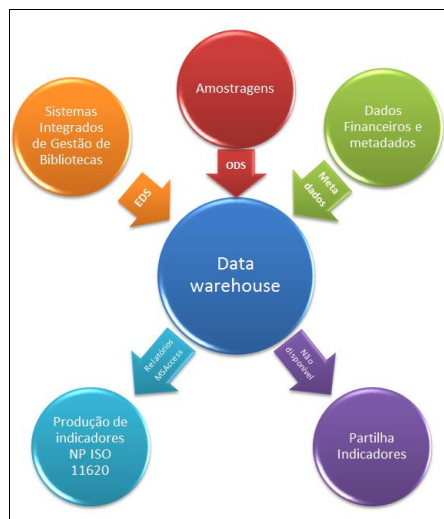


Figura 2.1.8: Estrutura do Inquidoc.

A Figura 2.1.9, em seguida apresentada diz respeito ao menu de indicadores do Inquidoc. Depois da extração dos dados e da sua transformação e carregamento na base de dados local *Access*, a geração dos indicadores é feita recorrendo a este menu, escolhendo as opções disponíveis nos botões.

Figura 2.1.9: Menu de indicadores do Inquidoc

A imagem apresentada na Figura 2.1.10 pretende demonstrar um relatório tipo, gerado pelo Inquidoc para os seus indicadores.

B.2.4.1 Movimentos da colecção de empréstimo	
Este indicador é calculado com base no número total de empréstimos e o número total de documentos numa colecção	
Formula de Cálculo	Dados
$\frac{A}{B}$	A 39076 B 144707
Indicador	0,27

gerado por: 17 de Dezembro de 2008

Figura 2.1.10: Exemplo de um relatório gerado pelo Inquidoc.

Em resumo, o Inquidoc representa um exemplo de uma aplicação, desenvolvida em contexto académico que emprega metodologias e práticas de armazéns de dados. Está contudo limitada à extração de indicadores ISO 11620 e à extração de dados de uma única fonte de dados, o Sistema de Gestão Integrada de Bibliotecas ALEPH.

2.2 Sistemas Comerciais

São apresentadas nesta secção algumas das soluções comerciais vocacionadas para o suporte à decisão que hoje em dia estão ao dispor dos Serviços de Documentação e Informação Universitários.

Todas elas são desenvolvidas por empresas líderes mundiais nas áreas do *software* de automação para Bibliotecas e no mercado da informação, contando cada uma delas com milhares de instituições na sua carteira de clientes.

2.2.1 Aleph Reporting Center

O *Aleph Reporting Center* (ARC) é um exemplo clássico de um armazém de dados, que permite interrogar, extrair e consultar relatórios e estatísticas sobre os dados do Sistema de Gestão Integrada de Bibliotecas ALEPH (ExLibris Group 2012, 6).

O ALEPH e o seu armazém de dados ARC são sistemas desenvolvidos e comercializados pela empresa *ExLibris Group*, líder mundial em sistemas de automação para Bibliotecas.

O Sistema de Gestão Integrada de Bibliotecas ALEPH constitui um dos sistemas mais difundidos e utilizados no mundo, sendo a solução adotada pela generalidade das Bibliotecas Universitárias Portuguesas, incluído a Biblioteca da Faculdade de Engenharia da Universidade do Porto.

O ARC representa um bom exemplo da aplicação prática de metodologias de armazéns de dados, com a particularidade de ser uma ferramenta dedicada ao tratamento de informação originária de um sistema operacional específico – O ALEPH. Dispõe de mais de 100 relatórios e estatísticas predefinidas, com a possibilidade de personalização (ExLibris Group 2007).

Os dados operacionais são recolhidos do ALEPH e apresentados aos utilizadores de uma forma lógica, apresentando uma estrutura fácil de usar e permitindo o cruzamento de dados entre

múltiplas áreas: *OPAC*, Catalogação, Circulação, Aquisições/Controlo de Periódicos, Empréstimo Interbibliotecário, Bibliografia de Disciplinas.

O ARC dispõe de um conjunto de ferramentas intuitivas que permitem criar relatórios e estatísticas, sem que os colaboradores da Biblioteca tenham a necessidade de ter conhecimentos de *SQL* ou da estrutura da base de dados (ExLibris Group 2007).

Esta solução é comercializada de forma integrada com o IBM *Cognos ReportNet™*, um *software* encarregue da camada lógica de apresentação, que permite gerir e gerar os relatórios e estatísticas (ExLibris Group 2012).

A Figura 2.2.1 apresenta uma visão conceptual do ARC.

Em contraste com a base de dados do ALEPH que é orientada para o registo e para trabalhar com transações, o armazém de dados do ARC é orientado para a extração rápida de dados, relatórios e estatísticas.

O ARC é constituído por três componentes principais (Shachar-Abutbul 2011):

- **Extração, Transformação e Carregamento (ETL)** – este componente extrai os dados das tabelas do ALEPH, transforma-os e coloca-os numa estrutura física diferente (esquema em estrela) e por fim carrega-os, já racionalizados no ARC;
- **Armazenamento** – este componente mantém e gere os dados transformados, recebidos da componente ETL;
- **Acesso** – este é o componente onde os utilizadores interrogam e analisam os dados armazenados.

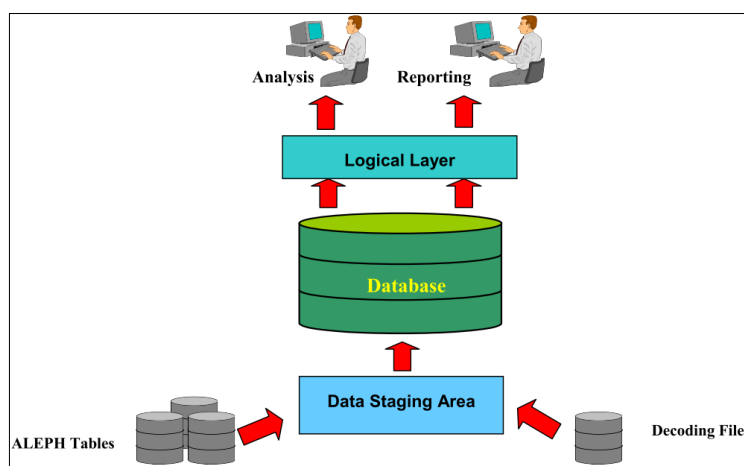


Figura 2.2.1: Visão conceptual do ARC. (ExLibris Group 2012, 5)

Estando as metodologias de armazéns de dados na génese do ARC, logicamente se compreende que o seu componente principal é o processo de extração, transformação e carregamento de dados (ETL).

O ARC é composto pelas seguintes camadas (Shachar-Abutbul 2011):

- **Mirroring área (MRR)** – A área onde é depositada uma imagem das tabelas extraídas do ALEPH. Esta camada é estática, isto é, as tabelas MRR não são processadas.
- **Staging área (STG)** – A área temporária para onde são extraídos os dados das tabelas MRR e onde são construídas tabelas com uma estrutura desnormalizada (esquema em estrela). A cada operação de carregamento de dados, esta área é reinicializada e carregada outra vez com dados. No fim desta etapa os dados das tabelas STG são carregadas na camada *End-user layer* (DWH).
- **End-user layer (DWH)** – A área do utilizador, que inclui tabelas com a mesma estrutura das tabelas STG. Esta camada é completamente indexada para suportar uma

pesquisa eficiente. Ligações entre tabelas serão definidas ao nível da ferramenta de relatórios (*Cognos ReportNet*).

A Figura 2.2.2 apresenta uma visão dos componentes e camadas do ARC.

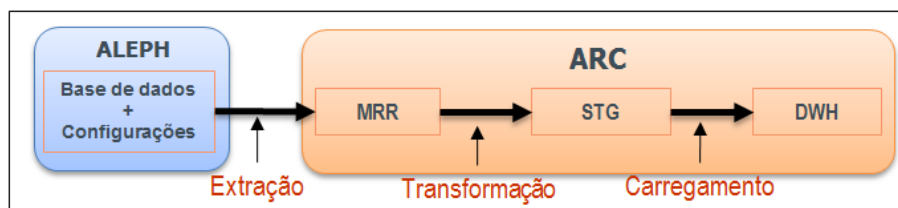


Figura 2.2.2: Processo ETL e camadas ARC.

Para terminar a caracterização do ARC, é apresentado em seguida um breve exemplo da potencialidade da aplicação. A Figura 2.2.3 ilustra a capacidade de extração de informação, recorrendo à geração automática de gráficos, listas e tabelas.

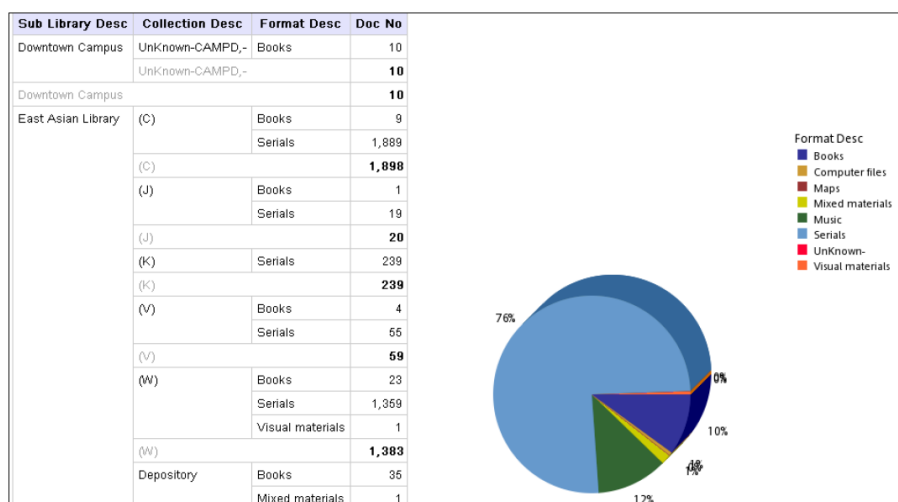


Figura 2.2.3: Exemplo de relatório gerado pelo ARC. (ExLibris Group 2012)

Sintetizando, o ARC constitui um bom exemplo da aplicação das metodologias de armazéns de dados, sendo no entanto limitado pelo facto de ser um armazém de dados dedicado a uma única fonte de dados, o Sistema de Gestão Bibliotecária ALEPH.

2.2.2 360

O 360 é um conjunto de produtos e serviços da *Serial Solutions* vocacionados para as atividades de suporte à decisão, cobrindo diversas áreas, desde controlo de coleções até à obtenção de estatísticas de utilização de recursos. Estas ferramentas e serviços podem ser adquiridas separadamente, de acordo com as necessidades de cada Biblioteca.

As ferramentas que constituem o pacote são (Serial Solutions 2012b):

- O *360 Core* – é a base para acesso, gestão e avaliação de serviços. É disponibilizado como SaaS. A Figura 2.2.4 representa o núcleo fundamental para quem quer adquirir e utilizar esta ferramenta;

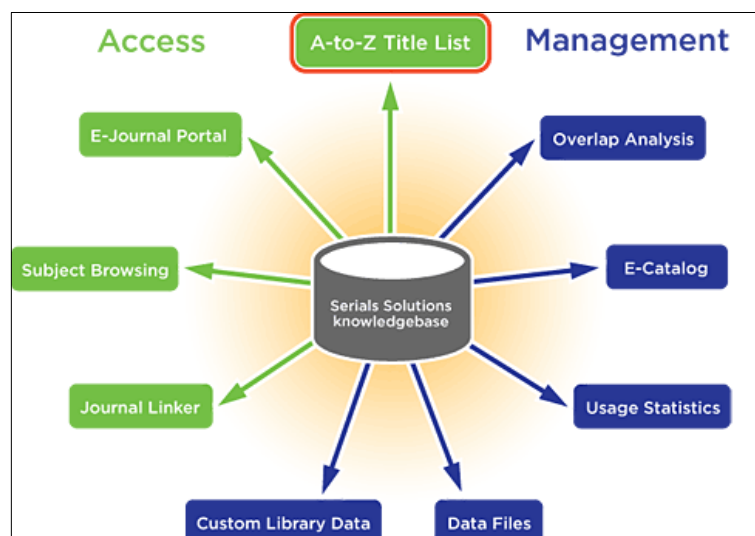


Figura 2.2.4: Núcleo fundamental para a utilização da solução 360 da Serials Solutions.

- O *360 Counter* – É uma ferramenta de agregação, avaliação e suporte à decisão, que elabora relatórios sobre estatísticas de uso das coleções das Bibliotecas, disponibilizadas pelos fornecedores. A Figura 2.2.5 representa um diagrama sobre o funcionamento do 360 Counter e ainda a comparação entre o esforço nos processos de avaliação, gestão e decisão, com e sem esta ferramenta, segundo a perspetiva do vendedor;

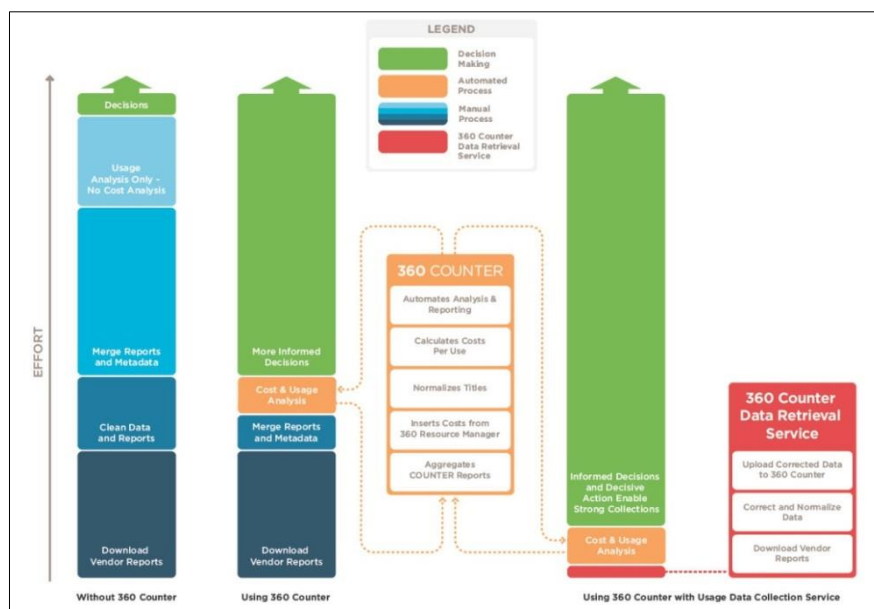


Figura 2.2.5: Esforço das tarefas de gestão, com e sem o 360 Counter, segundo a perspetiva do vendedor.

- O *360 Resource Manager* – É uma ferramenta para o controlo de custos, recursos e otimização de coleções. Disponibiliza vários relatórios: dados de subscrição e custos, licenciamento, termos e condições, informação administrativa, estatísticas de fornecedor, análises de reposição, estatísticas de utilização, entre outros. A Figura 2.2.6 apresenta uma visão geral da arquitetura dos produtos 360 da Serial Solutions.

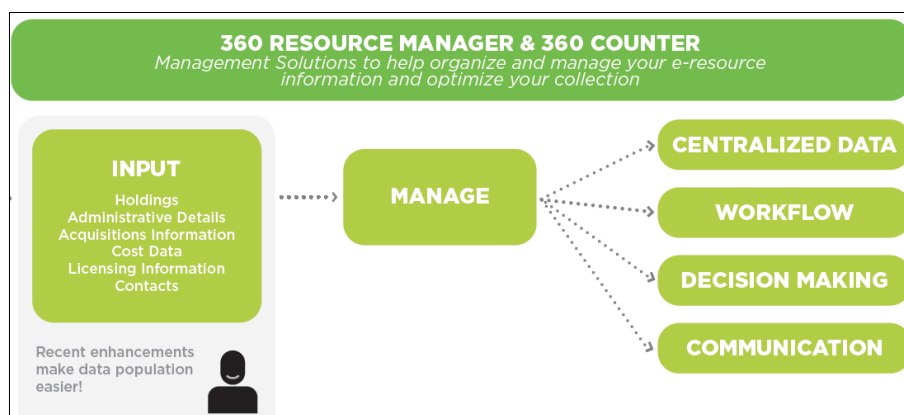


Figura 2.2.6: Arquitetura do produto 360 da Serial Solutions. (Serial Solutions 2012a)

Durante o seu trabalho de investigação onde leva a cabo um estudo comparativo de quatro sistemas de suporte à decisão apresentados neste documento (o *Counter 360*, o *Journal Use Reports*, o *ScholarlyStats* e o *Serials Analysis System*), Paynter (2008) tece as seguintes considerações sobre este produto: é um produto focado na análise de utilização de recursos, não permite o carregamento de relatórios de utilização das Bibliotecas mas faz uma verificação de qualidade aos relatórios da Biblioteca provenientes dos seus fornecedores, emitindo um alerta quando esses relatórios não estão de acordo com a norma COUNTER, permite análises custo/utilização e utilização/assunto, permite algum tipo de análise no próprio sistema.

Paynter (2008, 36,44) afirma que o 360 é a solução preferencial para Bibliotecas que estejam interessadas em análises custo/utilização e a preferida por Bibliotecas onde os recursos humanos são maiores porque, apesar de não permitir o carregamento de dados de utilização, tem mais funcionalidades e mais relatórios (comparativamente com a *ScholarlyStats*).

2.2.3 Journal Use Reports

O *Journal Use Reports* (JUR) é uma ferramenta de gestão e avaliação de coleções que permite combinar a informação da Thomson sobre citações científicas com os dados de utilização das Bibliotecas que subscrevam o acesso a este serviço. A combinação destas duas fontes de dados permite estabelecer um panorama sobre o estado da utilização dos recursos das Bibliotecas. O objetivo desta ferramenta é ajudar os gestores das Bibliotecas a priorizar e a avaliar a suas coleções, bem como ajudar os investigadores a demonstrarem o impacto e proeminência da sua investigação.

As origens de dados que alimentam esta ferramenta são (Thomson Reuters 2009):

- Dados do *Journal Citation Reports*, a base de dados sobre citações a Publicações Periódicas da Thomson Reuters;
- Dados da Web of Science, um índice de citações académicas da Thomson Reuters;
- Dados sobre as coleções da Instituição;
- Dados sobre a utilização das publicações periódicas subscritas pela instituição, fornecidas pelos fornecedores.

Em termos de relatórios suportados, esta ferramenta suporta o relatório COUNTER JR1 – Número de pedidos bem-sucedidos de acesso a texto integral de artigos, por mês e por revista (Paynter 2008, 37-38).

Ainda segundo Paynter (2008, 40), os tipos de análise suportados pelo JUR são a análise por periódico, por editor, por pacote de aquisição, por fontes de dados, por comparação a testes de referência e por análise do histórico de tendências.

Relativamente à apreciação global do produto, afirma que o JUR, a par do USAS, é o melhor na análise comparativa de coleções, detém alguma capacidade de análise nativa no sistema,

permite o descarregamento da informação para uma folha de trabalho Excel, não dispondo no entanto de funcionalidades administrativas que são hoje em dia consideradas normativas na área do suporte e não disponibiliza normalização de títulos, ISSN, eISSN (Paynter 2008, 34, 39, 43).

Apesar disso, Paynter (Paynter 2008, 44, 45), considera que o JUR pode ser apelativo para grandes Bibliotecas de investigação com colaboradores capazes de manipular os dados ou para Bibliotecas com fundos suficientes para adquirir os serviços do JUR a par dos serviços do 360 ou do *ScholarlyStat*. A vantagem comercial do JUR é a ligação a outros sistemas comerciais do produtor (*Journal Citation Reports* e *Web of Science*).

2.2.4 ScholarlyStats

Este sistema de suporte à decisão, desenvolvido pela Swets, providencia ferramentas para monitorar, manter e suportar a análise de estatísticas de utilização de recursos eletrónicos das Bibliotecas que subscrevam o serviço.

O *ScholarlyStats* encontra-se acessível através de um portal que constitui um ponto único de acesso para a apresentação das estatísticas de utilização, depois de ultrapassados os processos de recolha, normalização e consolidação dos dados.

Segundo a visão da Swets, existe uma proliferação de dados sobre a utilização dos recursos das Bibliotecas, proveniente de diversos fornecedores e que se apresentam de forma diversificada. A Figura 2.2.7 representa a formulação do problema e a respetiva solução, de acordo com o ponto de vista da Swets.

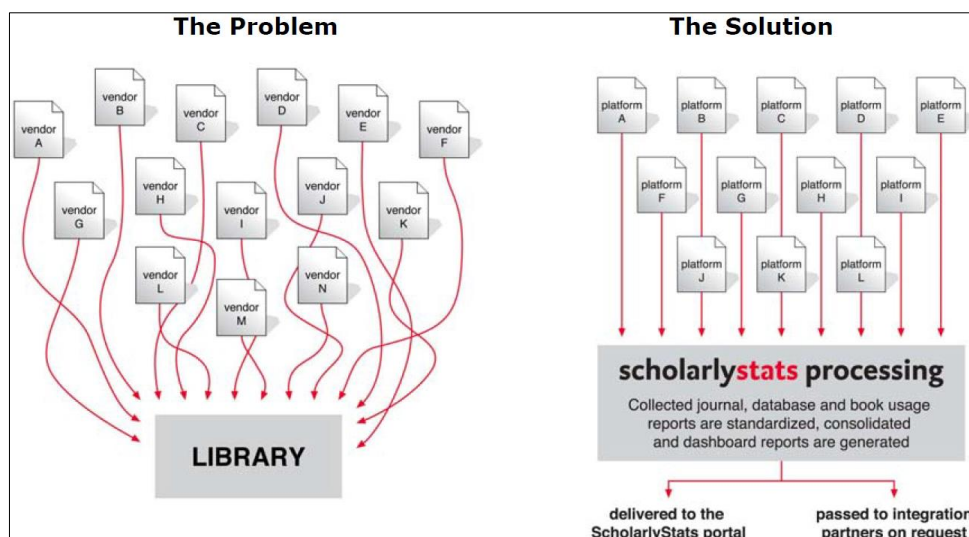


Figura 2.2.7: O problema e respetiva solução. A visão da Swets. (Swets 2008)

O processo de recolha de dados consiste no carregamento dos dados dos clientes, em formato COUNTER, sendo suportados os seguintes relatórios para recolha (Swets 2008):

- Número de pedidos de acesso bem-sucedidos a artigos em texto integral, por mês e por revista (relatório COUNTER JR1);
- Pesquisas e sessões por mês e por base de dados (relatório COUNTER DR 1);
- Acessos negados por mês e por base de dados (relatório COUNTER DR2);
- Pesquisas e sessões por mês e por serviço (relatório COUNTER DR3);
- Títulos pedidos por mês e por título do livro (relatório COUNTER BR3);
- Pedidos de secções por mês e por título do livro (relatório COUNTER BR2).

Após o processo de recolha, o *ScholarlyStats* normaliza e consolida os dados, transformando-os num formato privativo que facilita a análise das estatísticas de uso das revistas,

das bases de dados e dos livros eletrônicos ao longo de todas as plataformas, garantindo que os relatórios consolidados apresentam uma visão geral da utilização dos recursos, em colunas normalizadas, todas preenchidas com dados (Swets 2008).

Em termos de apresentação de resultados, o *ScholarlyStats* oferece um pacote de relatórios com estatísticas de uso em formatos compatíveis e não compatíveis com a norma COUNTER. Estes relatórios consolidados encontram-se disponíveis para os clientes no portal, com uma periodicidade mensal, nos formatos comprimido (zip), valores separados por vírgulas (csv) e Excel (xls).

Os tipos de relatórios disponíveis são (Swets 2008):

- Os Relatórios Consolidados *ScholarlyStats* – englobam toda a informação sobre estatísticas de uso de um cliente, consolidada numa única folha de cálculo;
- Os quadro-resumo *ScholarlyStats* – a informação resumida para uma rápida análise de uso e tendências;
- Relatórios originais – os relatórios originais, inicialmente carregados para análise;
- Arquivos de relatórios *ScholarlyStats* – um histórico dos relatórios gerados anteriormente, para permitir a análise de tendências ao longo dos meses e anos anteriores.

O *ScholarlyStats* suporta ainda a transferência de dados para outros sistemas, como por exemplo sistemas de gestão de recursos eletrônicos (ERM), através do protocolo SUSHI.

A Figura 2.2.8 apresenta três exemplos de tipos de relatórios disponíveis no *ScholarlyStats*.

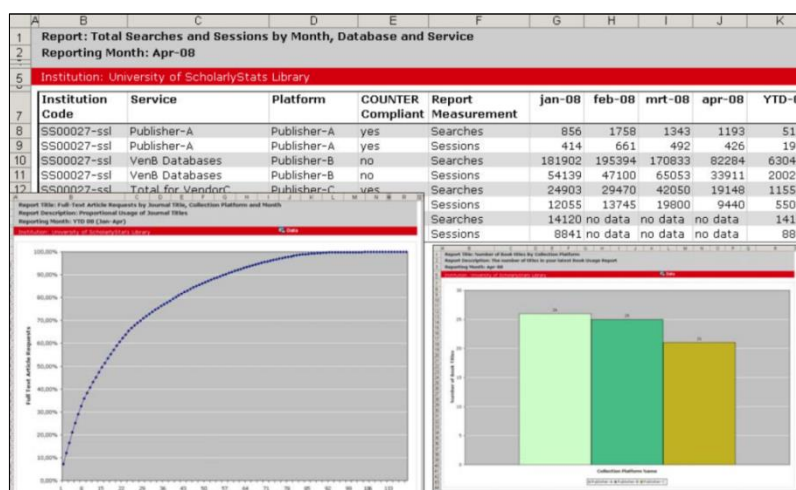


Figura 2.2.8: Relatórios *ScholarlyStats*.

Na sua análise, Paynter considera que muitas Bibliotecas se sentirão imediatamente atraídas pelo *ScholarlyStats* porque recolhe relatórios de utilização diretamente dos sítios web dos fornecedores das Bibliotecas, organizando-os e arquivando-os, resultando em economias significativas de tempo. Para além disto dispõe de funcionalidades interessantes, como sejam os quadros-resumo com macros para a análise e visualização automática de informação comumente procurada (Paynter 2008, 35).

2.2.5 Ulrich's Serials Analysis System

O *Ulrich's Serial Analysis System* (USAS) é um sistema de suporte à decisão, igualmente comercializado pela Serial Solutions, especializado na análise e elaboração de relatórios de estatísticas de uso e manutenção de coleções de publicações periódicas.

Permite estabelecer testes de comparação entre a coleção da instituição e outras coleções de outras instituições ou grupos de instituições com características demográficas semelhantes, testes

de comparação à coleção da instituição que está integrada num consórcio e ainda testes comparativos entre a coleção da instituição e coleções de publicações periódicas pertencentes à base de conhecimento da Ulrich (Serial Solutions 2012c; Paynter 2008, 34).

Permite ganhar conhecimento sobre a profundidade, abrangência, pontos fortes e pontos fracos das coleções em formato impresso ou eletrónico, bases de dados de texto integral e de fontes de resumos e indexação (Serial Solutions 2012c).

Disponibiliza relatórios que fornecem análises detalhadas, título a título, de revistas impressas, revistas eletrónicas e bases de dados de publicações periódicas, disponíveis através de *download*, por exemplo em formato Excel, trabalhados sempre fora do SSD (Paynter 2008, 39).

Segundo Paynter (2008), o USAS é, a par com o JUR, a solução que melhor faz análises comparativas de coleção, aproveitando o saber histórico acumulado da empresa. É também o sistema de suporte à decisão recomendado por Paynter (2008, 36) para comparar preços de aquisição de revistas englobadas em pacotes versus preços de revistas individuais, muito devido à posição histórica detida pela empresa no mercado editorial de periódicos. Por fim Paynter (2008, 45), afirma ainda que a força do *Ulrich's Serial Analysis System* reside na capacidade que esta ferramenta tem em fazer análises de alto nível, que incluem análises de custos por utilização, análise comparativa de coleções, análises de pacotes de revistas e partilha de recursos.

O *Ulrich's Serial Analysis System* assenta todas as suas potencialidades de análise e suporte à decisão na *Ulrich's knowledgebase*. É uma base de dados com mais de 300.000 publicações periódicas, de 90.000 editores, cobrindo 950 áreas, em 200 línguas (Serial Solutions 2012d).

2.3 Tecnologia

Em seguida são apresentados alguns exemplos das infraestruturas de suporte, modelos de implementação e ferramentas disponíveis para o desenvolvimento de armazéns de dados e respetivas soluções de apoio à decisão.

2.3.1 Infraestruturas de suporte à implementação de armazéns de dados

São várias as soluções disponíveis, comerciais ou gratuitas, para suportar a implementação de um armazém de dados.

Nesta secção elencam-se algumas das mais populares soluções disponíveis, desde as ofertas de marcas comerciais como a ORACLE® ou a Microsoft®, até às propostas provenientes da comunidade de desenvolvimento de *software* livre, como por exemplo *PostgreSQL* ou *MySQL*.

ORACLE (LUMPKIN 2009)

A ORACLE é uma empresa líder mundial em soluções de gestão de dados, oferecendo uma plataforma abrangente para armazéns de dados, baseado no seu sistema de gestão Oracle 11g, em tecnologia líder de mercado nas áreas de desempenho e escalabilidade, na integração de ferramentas de análise de dados e na incorporação de mecanismos de verificação e manutenção da qualidade dos dados.

Para além da base dados ORACLE 11g, a solução possui também várias ferramentas analíticas, integrando OLAP, mineração de dados e funcionalidades estatísticas diretamente suportadas na base de dados.

Seguindo os requisitos de integração de qualquer armazém de dados, a solução da ORACLE inclui também o *Oracle Warehouse Builder* (OWB)², uma ferramenta de implementação de armazéns de dados que permite a extração, transformação e carregamento de dados, utilizando as capacidades de transformação escalável e o acesso a dados heterogéneos da ORACLE.

² Para mais informações ver: Ferramentas de suporte à modelação de armazéns de dados

Apresentam-se em seguida algumas das características mais importantes para a construção de um armazém de dados, oferecidas pela Oracle:

- Desempenho
 - Conjunto alargado de técnicas de otimização:
 - Técnicas de indexação e agregação avançadas;
 - Otimização de interrogações e vistas materializadas baseadas em OLAP para interrogações dimensionais;
 - Métodos de acesso a interrogações escaláveis e paralelizadas para o processamento de grandes volumes de dados.
 - Um gestor de recursos para assegurar desempenho em armazéns de dados complexos;
 - Um modelo único e patenteado de leitura consistente, para assegurar que não existe impacto do carregamento de dados no desempenho das interrogações;
 - Pacotes de gestão para automatizar a operação de um armazém de dados;
 - Execução paralela em memória.
- Escalabilidade
 - Particionamento;
 - Compressão;
 - Agrupamento;
 - Paralelismo;
- Análise
 - OLAP;
 - Mineração de dados;
 - Estatísticas.
- Integração
 - OWB
 - ODI
 - Motores de análise (OLAP, mineração de dados e estatísticas)

MICROSOFT (2012B)

A integração entre a base de dados *SQL Server 2008 R2* e a sua plataforma de inteligência organizacional, formam a oferta da Microsoft na área de gestão e armazenamento de dados. A Figura 2.3.1 apresenta uma visão da arquitetura proposta pela Microsoft para a implementação de um armazém de dados.

À semelhança da Oracle, a oferta da Microsoft apresenta a integração de uma vasta gama de ferramentas que facilitam a implementação de armazéns de dados e a sua análise, contando também com as mais avançadas características atualmente disponíveis e que garantem à sua solução boas prestações de desempenho e escalabilidade.

Algumas das principais características são:

- Compressão de dados;
- Requisitos de armazenamento reduzidos e alto desempenho;
- Gestor de recursos de cargas de trabalho concorrentes;
- Declaração *Transact-SQL MERGE* para realizar tarefas comuns no armazém de dados;
- Paralelismo e otimização de interrogações com esquema em estrela de forma a acelerar as consultas ao armazém de dados;
- Registo de log reduzidos relativos às operações *INSERT INTO* e *SELECT FROM*, com o objetivo de acelerar os processos de extração, transformação e carregamento;
- Tarefas de pesquisa escaláveis no *SQL Server Integration Services*, com o objetivo de aumentar o desempenho do processo ETL;
- Identificação de alterações e captura de dados em base de dados operacionais.

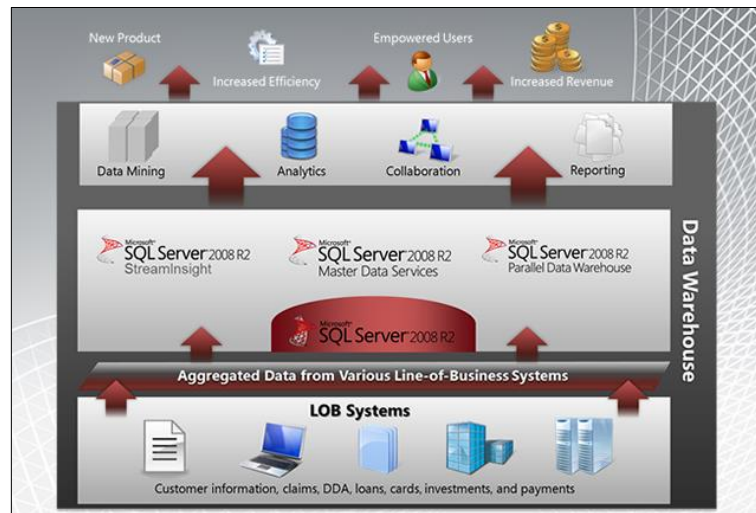


Figura 2.3.1: Arquitetura de um armazém de dados da Microsoft (Microsoft Corporation 2012a)

As operações de ETL são implementadas recorrendo ao *SQL Server Integration Services* (SSIS). Elencam-se em seguida algumas das suas potencialidades:

- Arquitetura de linha de montagem escalável, multitarefa e de 64bits;
- Conectividade a servidores *noSQL*, *RDBMS*, sistemas *mainframe* e sistemas ERP;
- Elevado número de transformações pré-compiladas para consolidação de dados de vários sistemas, redução de duplicação e limpeza de dados;
- Capacidade para gerir dimensões de alteração lenta (*slowly changing dimensions*);
- A integração com o *Analysis Services* do *SQL Server* para a construção direta de cubos;
- Extensibilidade através da *Microsoft.NET Framework*, incorporando *scripts* personalizados diretamente no fluxo de integração de dados;
- Integração de elevado volume de dados através da consulta a tabelas de grandes dimensões;
- Utilização da funcionalidade de identificação e captura de dados em sistemas operacionais para evitar processamento em lote de grande volume de dados. Aproveitamento da nova instrução *Transact-SQL MERGE*.

No campo das ferramentas de análise para a exploração de dados, a solução de gestão e armazenamento de dados da Microsoft conta com um motor analítico escalável, capaz de manipular grandes conjuntos de dados. Apresenta os tradicionais relatórios, com análise OLAP, com mineração de dados e ainda com indicadores chave de desempenho (KPI) e painéis de resultados (*scorecards*), tudo isto num modelo dimensional unificado (UDM) que inclui todas as entidades e medidas do armazém de dados.

No campo da visualização e colaboração, o armazém de dados da Microsoft fornece análises às aplicações que os decisores utilizam diariamente, recorrendo para isso ao *Microsoft Office PerformancePoint® Server* como forma de integração com os outros produtos de inteligência organizacional da Microsoft, como sejam, o *Microsoft Office System 2010*, o *SQL Server 2008 R2 Reporting Services* e o *Microsoft Office SharePoint® Server 2010*.

POSTGRESQL

A PostgreSQL é uma base de dados de código aberto, contando com mais de 15 anos de contributos de uma numerosa comunidade de desenvolvedores.

Suporta os principais sistemas operativos, como por exemplo, Linux, Unix, Mac e Windows.

Ganhou notoriedade devido à sua arquitetura fiável e à sua consistência de dados, garantindo as propriedades ACID (Atomicidade, Consistência, Isolamento, Durabilidade) e suportando as principais linguagens de programação PERL, PHP, C/C++, Java, C#, Python, entre outras.

Alguns dos pontos fortes desta base de dados são (PostgreSQL; Bartolini 2009; Berkus 2011):

- O suporte SQL, incluindo cursores, funções, procedimentos armazenados, interrogações preparadas, integridade referencial, subinterrogações, transações e sub-transações, gatilhos e vistas;
- Boa escalabilidade;
- Bom desempenho;
- Alta disponibilidade;
- Replicação;
- Cópias de segurança em linha;
- Ambientes de desenvolvimento gráficos;
- Ferramentas de análise e gestão;
- Cometimento (*commit*) em duas fases;
- Extensibilidade, incluindo tipos de dados e domínios definidos pelo utilizador;
- Particionamento horizontal suportado através de herança;
- Funcionalidades de *Windowing*;
- Interrogações com cláusula *WITH*.

É uma base de dados que pode ser usada na implementação de um armazém de dados, apesar de carecer de algumas características presentes nas soluções de armazenamento de dados comerciais analisadas anteriormente.

Alguns dos pontos fracos são (Bartolini 2009; EnterpriseDB 2010):

- Falta de suporte nativo para distribuição de dados e processamento paralelo (carência de uma cláusula “*DISTRIBUTED BY*”);
- Falta de suporte para índices bitmap só em memória. Suporte só para índices em disco;
- Falta de suporte transparente para particionamento de dados (carência de uma clausula “*PARTITION BY*”);
- Falta de suporte transparente para vistas materializadas, atualmente simuladas recorrendo à combinação de procedimentos armazenados e vistas;
- Falta de suporte melhorado para extensões temporais. Tipo de dados “*period*”, funções de comparação temporal (*precedes, overlaps, contains, meets*).

Outra diferença notada relativamente às soluções comerciais estudadas é a inexistência de uma plataforma integrada de ferramentas que abranja os processos de modelação, de ETL e de análise e inteligência organizacional. Dito por outras palavras, o seu formato de apresentação resume-se à base de dados, apesar de suportar integração com outras ferramentas, como por exemplo o *Kettle* para o desenvolvimento dos processos de ETL, o *Pentaho Business Analytics* para a exploração de dados e inteligência organizacional ou ainda o *dblink*, o *PL/Proxy* ou o *DBI-Link* para a ligação a fontes de dados externas (Bartolini 2009).

MySQL

A base de dados *MySQL* é outra opção credível para a implementação de um armazém de dados, criada pela *MySQL AB*, uma empresa que operava no domínio do *software* livre, adquirida pela *Sun Microsystems* em 2008, que por sua vez foi adquirida em 2010 pela *Oracle*.

“A sua velocidade e fiabilidade tornaram-na uma alternativa popular aos sistemas de bases de dados comerciais”(Di Giacomo 2005).

Atualmente, o MySQL é disponibilizado numa versão gratuita, a *MySQL Community Edition*, e quatro versões pagas, onde a composição dos pacotes disponibilizados varia de acordo com os requisitos de negócio dos clientes (MySQL 2011a).

Algumas das características que a tornam apropriada para a implementação de um armazém de dados são (MySQL 2012a):

- Particionamento de dados/índices (intervalo, dispersão, chave, lista, composto);
- Praticamente sem limites de armazenamento (1 partição de tabelas = 110TB) com gestão de armazenamento automático;
- Replicação nativa;
- Forte suporte à indexação (árvores B, texto integral, agrupado, dispersão, GIS);
- Cache de dados e índices, múltiplos e configuráveis;
- Carregamento prévio de dados em cache;
- Cache de interrogação único (não guarda só os dados, guarda resultado e interrogação);
- Paralelismo no carregamento de dados;
- Linguagem de manipulação de dados multi-inserção;
- Tabelas só para leitura (*Read-only*);
- Optimizador baseado no custo;
- Vasta plataforma de suporte.

Suporta ainda diversos motores de armazenamento de dados nativos: *InnoDB*, *MyISAM*, *Cluster*, *Federated*, *Archive*, *Merge*, *Memory*, *CSV* e *Blackhole* (MySQL 2012a).

O núcleo da base de dados pode ser complementado com uma vasta gama de ferramentas que facilitam a construção e gestão de um armazém de dados (MySQL 2011a, 2011b):

- *MySQL Enterprise Backup* – Ferramenta que permite salvarguardar os dados. Suporta salvaguarda de dados realizada com a base de dados em funcionamento, salvaguarda de dados incremental, salvaguarda de dados parciais. Suporta também algoritmos de compressão de dados;
- *MySQL Enterprise High Availability* – Permite a alta disponibilidade da infraestrutura, através de soluções de suporte e certificação;
- *MySQL Enterprise Scalability* – permite suportar um desempenho sustentado e requisitos de desempenho crescentes, através do *MySQL Thread Pool*, um gestor de processos desenhado para reduzir a sobrecarga das conexões de clientes e da execução de interrogações;
- *MySQL Enterprise Security* – disponibiliza módulos de autenticação externa prontos a utilizar e de fácil integração com as infraestruturas de segurança existentes, incluindo módulos de autenticação Linux (PAM), serviços de diretório Windows (*Active Directory*) e LDAP;
- *MySQL Enterprise Monitor* - O *MySQL Enterprise Monitor* e *MySQL Query Analyzer* fornecem uma visão do estado de saúde das bases de dados MySQL e emitem alertas para potenciais problemas que poderão afetar o sistema;
- *MySQL Workbench* – é uma ferramenta visual e unificada que permite aos desenvolvedores, administradores de bases de dados e arquitetos de dados

desenharem, desenvolverem e administrarem as referidas bases. Possui avançadas técnicas de modelação, um editor SQL e ferramentas de administração que permitem o desenho, desenvolvimento de interrogações e a configuração e administração do servidor de base de dados.

À semelhança do *PostgreSQL*, a solução oferecida pelo *MySQL* não dispõe de ferramentas para ETL nem para exploração e análise de dados, sendo necessário recorrer a *software* de terceiros. O *Pentaho Business Analytics*, já referido para o caso do *PostgreSQL*, permite superar essas lacunas, quer ao nível do desenho dos processos de extração e transformação de dados (através do *Kettle*), quer ao nível da análise e exploração de dados (Pentaho 2012).

Uma versão completa das especificações técnicas poderá ser consultada no anexo A1.

Um quadro comparativo de algumas características das quatro bases de dados apresentadas neste estudo pode ser consultado no anexo A2.

2.3.2 Ferramentas de suporte à modelação de armazéns de dados

O ORACLE WAREHOUSE BUILDER E O ORACLE DATA INTEGRATOR (LUMPKIN 2009)

O *Oracle Warehouse Builder* é uma ferramenta de integração de dados que permite a construção de armazéns de dados de qualquer dimensão. Inclui um repositório de metadados multiutilizador, capacidade de modelação de dados e uma vasta variedade de técnicas de extração e transformação, para além da escalabilidade e desempenho de uma arquitetura ETL. É incluído sem custos adicionais associados na base de dados ORACLE 11g.

As características avançadas do OWB, anteriormente parte da opção empresarial ETL da Oracle, foram fundidas no *Oracle Data Integrator Enterprise Edition*. O *Oracle Data Integrator* (ODI) oferece não só as características ETL, frequentemente necessárias nos armazéns de dados, como também oferece uma plataforma de integração para, por exemplo, serviços de dados SOA e fluxos de dados baseados em eventos (*event-driven data feeds*).

Atualmente a Oracle inclui as características avançadas do ODI e do OWB como duas componentes do *Oracle Data Integrator Enterprise Edition*. Futuramente estes produtos fundir-se-ão numa tecnologia de integração de dados unificada. Esta tendência preserva no entanto qualquer investimento de desenvolvimento dos clientes da Oracle e providenciará mecanismos de atualização dos atuais componentes para a plataforma unificada.

Em vez de oferecer um motor de transformação de dados externo, o OWB usa uma arquitetura ETL, executando as suas transformações dentro da base de dados Oracle, aproveitando todo o seu desempenho e escalabilidade. A base de dados Oracle tem um conjunto abrangente de ferramentas de extração, transformação e carregamento. Estas características são alavancadas pelo OWB mas podem também ser usadas em qualquer ambiente onde dados são integrados num armazém de dados Oracle. Algumas dessas características incluem:

- Funcionalidades para acesso a sistemas não-Oracle;
- Utilitário de carregamento para carregamento expedito de ficheiros;
- Extensões SQL para transformação de dados;
- Funções de tabela: transformações definidas pelo utilizador, paralelas e eficientes;
- Captura de dados baseada em *logs* e baixa latência.

KETTLE (PENTAHO COMMUNITY 2012)

O Kettle faz parte de um projeto de *software* livre conhecido como *Pentaho Data Integration Community Edition* (PDI CE) e tem como objetivo colmatar as lacunas existentes nas áreas de extração, transformação e carregamento de dados, normalmente associadas a projetos de construção de armazéns de dados nas quais a solução envolve bases de dados

sem suporte adequado, como é o caso das bases de dados *PostgreSQL* e *MySQL* anteriormente estudadas.

Possui um ambiente gráfico e intuitivo (Figura 2.3.2) e uma arquitetura escalável e baseada em normas.

Suporta um elevado número de formatos de entrada e saída, onde se incluem ficheiros de texto, folhas de dados (tipo Excel) e ainda motores de bases de dados comerciais e gratuitas.

Cada processo é criado com uma ferramenta gráfica onde as ações são especificadas sem escrever código. Devido a esta característica, o *Kettle* é classificado como uma ferramenta orientada a metadados.

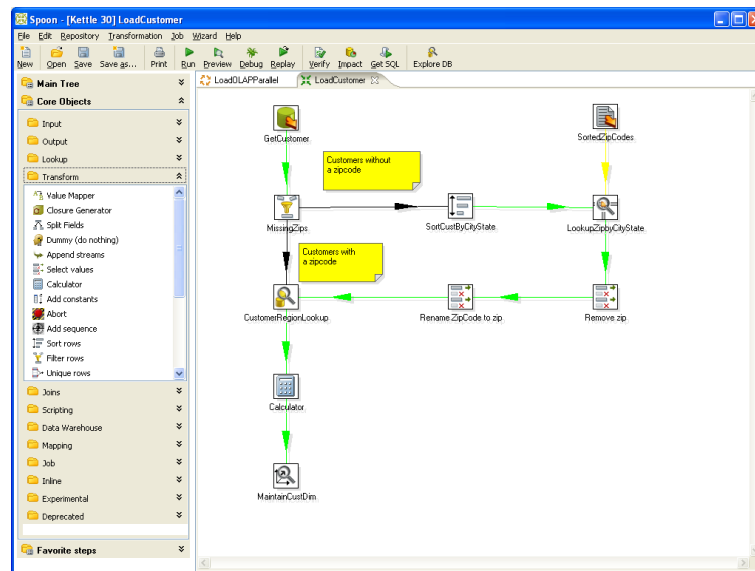


Figura 2.3.2: Desenho de uma transformação em *Kettle* (Pentaho 2008)

IBM COGNOS BUSINESS INTELLIGENCE (EXLIBRIS GROUP 2012; IBM 2012)

O *IBM Cognos Business Intelligence* é o sucessor do *Cognos ReportNet*, a aplicação adotada no já aqui referido ARC para a camada de apresentação ao utilizador. Este tipo de aplicações insere-se no domínio da inteligência organizacional, permitindo a análise e exploração de dados.

Genericamente esta solução disponibiliza:

- Um interface web que permite a interação entre os utilizadores e as fontes de dados através de relatórios previamente construídos;
- Uma ferramenta para gerar relatórios *ad hoc* simples, utilizando o conceito de arrastamento de objetos (*drag-and-drop*) para uma janela de construção de relatórios, criando assim uma camada de abstração que permite que os utilizadores não necessitem de saber SQL. A Figura 2.3.3 é um exemplo desta funcionalidade, neste caso o *Query Studio* do *IBM Cognos Business Intelligence* integrado com o armazém de dados do ARC;
- Uma ferramenta profissional para construir relatórios mais complexos e sofisticados que analisa os dados de acordo com as necessidades dos utilizadores. Oferece muitos métodos para formatar, apresentar e distribuir os relatórios gerados.

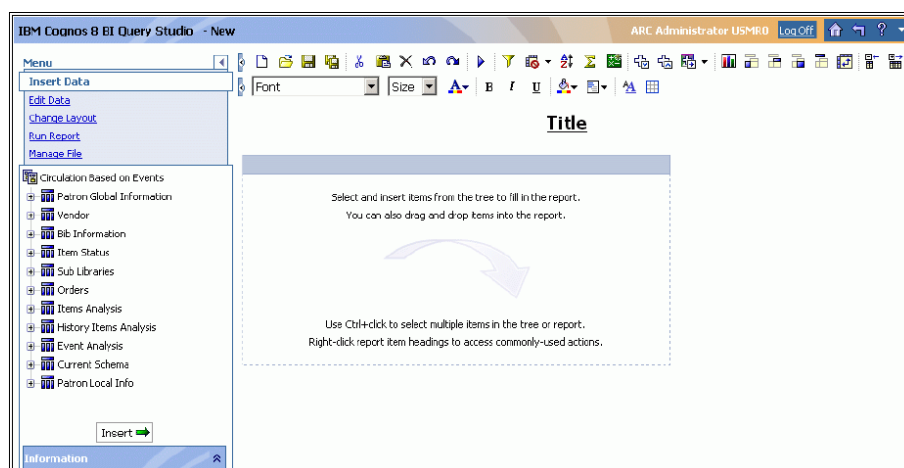


Figura 2.3.3: Query Studio do Cognos BI, integrado com o armazém de dados do ARC. (ExLibris Group 2012)

2.4 Indicadores de medida

Nesta secção apresentam-se alguns exemplos de indicadores de medida, especificados e desenvolvidos para o domínio dos Serviços de Documentação e Informação.

Estes instrumentos de medida poderão ser reaproveitados e implementados no armazém de dados ao mesmo tempo que poderão servir de referência para derivar novos indicadores de medida, ajustados a necessidades específicas do Serviço, principalmente para as áreas com carências normativas.

2.4.1 Indicadores para Bibliotecas

NORMA ISO 11620 (IPQ 2004, 2006; NEVES 2007)

A Norma Portuguesa ISO 11602 é idêntica à sua congénere internacional ISO 11620:1998, traduzida para Português e aprovada em 2004. O seu campo de aplicação incide na produção de indicadores de medida para bibliotecas de todo o tipo, de todos os países. Estes indicadores servem para efetuar comparações, ao longo do tempo numa mesma biblioteca ou entre bibliotecas equivalentes. Os indicadores são elementos de avaliação da qualidade e eficácia dos serviços de uma biblioteca, assim como indicadores da utilização eficiente dos recursos de que dispõe para atingir os seus objetivos. Esta norma não comporta indicadores relativos à avaliação do impacto das bibliotecas, quer ao nível dos indivíduos, quer ao nível da sociedade.

A análise da norma (IPQ:2004) permite verificar a extração de indicadores de três tipos. O primeiro tipo (B.1, consultar anexo B.1) está diretamente ligado aos clientes da biblioteca e à sua perceção em termos de qualidade e satisfação com os serviços disponibilizados pela organização. Este tipo de indicador pode ser obtido, tal como indicado na norma, com base em questionários que possibilitem a recolha desta informação junto do público, de forma a aferir a satisfação dos utilizadores.

O segundo tipo (B.2, consultar anexo B.1) está ligado aos serviços públicos disponibilizados pela biblioteca. Este tipo de indicador é gerado com base em operações de medição claras e bem definidas. Permitem aferir o desempenho dos serviços prestados pela biblioteca. A maior parte destes indicadores é obtida com base numa medição periódica dos serviços públicos disponibilizados aos utilizadores. São indicadores independentes, objetivos e de fácil extração dos sistemas de informação disponíveis nas bibliotecas.

O terceiro tipo de indicadores (B.3, consultar anexo B.1) permite efetuar uma autoavaliação aos serviços internos da biblioteca. Os indicadores deste tipo destinam-se a evidenciar o desempenho dos serviços internos das Bibliotecas, podendo ser utilizados para comparações entre bibliotecas.

Por último são indicados na norma (na sua adenda de 2006) um grupo (B.5) que permite obter os indicadores relacionados com a disponibilização e utilização dos recursos humanos. Estes indicadores são aplicados na avaliação da utilização e rentabilização dos recursos humanos disponíveis pela organização na prestação de serviços de atendimento público.

Uma lista destes indicadores pode ser consultada no anexo B.1

ISO 2789 (ISO 2003)

A norma ISO 2789 fornece recomendações para bibliotecas e a comunidade de serviços de informação sobre a extração e divulgação de estatísticas.

Os seus dois principais objetivos são estabelecer uma taxonomia dos elementos constituintes de um serviço de biblioteca para fins estatísticos e recomendar como cada um desses elementos deve ser medido. Além dos propósitos originais, tem também como objetivo especificar os tipos de dados requeridos pela norma ISO11620.

2.4.2 Relatórios Counter

Lançado em Março de 2002, o projeto COUNTER (*Counting Online Usage of Networked Electronic Resources*) é uma iniciativa internacional ao serviço das bibliotecas, editores e intermediários que visa o estabelecimento de normas que facilitem o registo e informação de estatísticas de utilização de recursos consultados na *internet*, de forma consistente, credível e compatível (Project COUNTER 2012).

É uma iniciativa que vem ao encontro das convicções convergentes de produtores e consumidores de informação, que defendem a necessidade de medir a utilização desses recursos de forma mais rigorosa. Os bibliotecários querem entender melhor de que forma a informação adquirida está a ser utilizada. Os editores dessa informação estão interessados em entender como é que os seus produtos disseminados estão a ser acedidos (Project COUNTER 2012).

O projeto COUNTER tenta responder a estas expectativas, através dos seus códigos de boas práticas, que mais não são do que um conjunto de normas e protocolos que ditam regras para o registo e a troca de dados de utilização (Project COUNTER 2012).

Atualmente estes são os códigos de boas práticas disponíveis (Project COUNTER 2012):

- 4.^a Edição do código de boas práticas COUNTER para Recursos Eletrónicos, publicado em Abril de 2012;
- 3.^a Edição do código de boas práticas COUNTER para Revistas e Bases de Dados, publicado em Agosto de 2008;
- 1.^a Edição do código de boas práticas COUNTER para Monografias e Trabalhos de referência, publicado em Março de 2006.

Um conjunto de listas dos relatórios sugeridos por estes códigos de boas práticas pode ser consultado no anexo B.2

2.5 Conclusões

Apresentam-se em seguida as principais ilações resultantes do trabalho de investigação levado a cabo na revisão da literatura.

Quanto aos trabalhos de investigação analisados, verificou-se uma carência considerável na discussão dos modelos de dados mais apropriados para a implementação dos armazéns de dados

dos projetos onde estes se inseriam. Exceção feita ao trabalho de investigação realizada sobre *Bibliomining*, a forma como os dados são dispostos e guardados num armazém de dados quase nunca chega a ser debatida. Outro ponto que ficou evidenciado na análise realizada às investigações consideradas relevantes, foi a carência ao nível de fontes de dados de suporte a um dado armazém de dados e a carência ao nível da abrangência/cobertura que um dado armazém de dados oferece à estrutura organizacional que serve. De facto verificou-se que estes trabalhos eram em grande medida muito centrados na área das bibliotecas, quando um Serviço de Documentação e Informação poderá oferecer outras valências, como por exemplo um museu, um arquivo ou até mesmo serviços editoriais. Dentro da área das bibliotecas, o foco da investigação foi sempre muito orientado, tanto quanto se pode inferir da literatura consultada, na deteção do acesso e utilização de recursos de informação. Não que esta área não seja importante, mas também existem áreas igualmente importantes que merecem ser cobertas e que são determinantes no processo de decisão de um Serviço de Documentação e Informação Universitário.

O estabelecimento de um modelo dimensional global para um Serviço de Documentação e Informação Universitário permitirá colmatar duas carências aqui identificadas, propondo à discussão um modelo de dados, com a vantagem de este ser global e abrangente de toda a organização.

Relativamente aos sistemas de apoio à decisão comerciais analisados, são sistemas do tipo “caixa negra”, verificando-se por isso o mesmo fenómeno identificado no caso dos trabalhos de investigação: carências ao nível da discussão do modelo de dados. Outro ponto menos positivo identificado foi mais uma vez o nível de abrangência oferecido por estas soluções, havendo inclusive estudos que concluem que, por si só, nenhuma das soluções apresentadas oferece uma taxa de cobertura total e satisfatória (Paynter 2008).

No caso do *Aleph Reporting Center*, a solução peca por se restringir a uma única fonte de dados, tornando-a pouco abrangente. A implementação de um armazém de dados à medida fará com que a organização beneficie de características importantes, como por exemplo a flexibilidade e a cobertura de múltiplas facetas da sua atividade.

Relativamente às tecnologias analisadas, verificou-se que qualquer uma das infraestruturas de suporte estudadas é credível para a implementação de armazéns de dados. As soluções comerciais têm como desvantagem os custos acrescidos. Por outro lado oferecem um nível de integração de ferramentas de suporte à implementação superior, quando comparadas com as soluções não comerciais, contemplando todo o ciclo de vida de um armazém de dados (modelação, implementação, ETL, análise e exploração de dados).

Nas soluções de *software* livre a vantagem é o reduzido custo associado. Por outro lado, o nível de integração de ferramentas é inferior, havendo necessidade de recorrer a aplicações de terceiros para colmatar essas carências.

No que toca ao domínio da especificação de indicadores de medida, a investigação realizada verificou que a área das bibliotecas é a grande impulsionadora no estabelecimento de padrões de avaliação e medida, tendo sido notadas carências a este nível na área arquivística, que recorre frequentemente à área das bibliotecas para a extrapolação dos seus indicadores. Na área museológica a carência de padrões foi também registada.

Pelos motivos acima referidos, acredita-se que a concretização deste trabalho de investigação e a implementação de um armazém de dados, em resposta ao caso de estudo concreto do Serviço de Documentação e Informação da Faculdade de Engenharia da Universidade do Porto, constitui uma mais-valia que vem dar resposta às carências evidenciadas, quer ao nível do SDI da FEUP, quer ao nível das soluções apresentadas neste capítulo.

3 Metodologia

Neste capítulo descrevem-se pormenorizadamente todas as etapas que servem de suporte ao desenvolvimento da dissertação, de acordo com os objetivos gerais e específicos definidos na introdução e com o enquadramento teórico apresentado no capítulo 2.

3.1 Metodologia de investigação e implementação

Tendo presentes os objetivos propostos, que sinteticamente se traduzem na obtenção de indicadores de medida fiáveis e de qualidade, recorrendo ao desenvolvimento de um armazém de dados, foi efetuada, em primeiro lugar, a revisão bibliográfica apresentada no capítulo anterior.

Essa revisão permitiu a fundamentação teórica e o levantamento do estado da arte em termos de conceitos e questões relacionadas com sistemas de suporte à decisão ao dispor de Serviços de Documentação Universitários, tendo sempre como pano de fundo conceitos, modelos e tecnologias associadas à implementação de armazéns de dados.

Perante a natureza da presente investigação, esta seguirá uma abordagem de estudo de caso, assente no entanto em metodologias próprias de um projeto de Engenharia Informática, quer na extração dos requisitos, quer no processo de implementação do mesmo.

Neste contexto, a investigação apresentará na sua fase de extração de requisitos um modelo em cascata, com várias interações com as diversas partes interessadas (*stakeholders*), e numa segunda fase apresentará uma metodologia de desenvolvimento iterativo, assente num processo cíclico de desenvolvimento de cada uma das estrelas do armazém de dados.

A Figura 3.1.1 apresenta uma visão geral da metodologia a adotar, seguindo-se uma descrição detalhada de cada fase do projeto, bem como os resultados esperados.

O processo inicia-se com uma reunião com a Direção do Serviço de Documentação e Informação com o intuito de clarificar os objetivos do projeto.

Segue-se uma fase de análise organizacional, com vista ao levantamento dos processos envolvidos e à determinação dos seus responsáveis. O desenvolvimento desta atividade será concretizado através da análise dos documentos identitários e regulamentos da organização, com a devida validação dos seus dirigentes intermédios. O produto final desta fase materializar-se-á no desenho dos casos de uso associados a cada processo.

Na fase seguinte serão identificadas as fontes de dados onde as atividades identificadas na fase anterior são de alguma forma registadas. No final desta fase será construída uma matriz que cruzará os processos com as fontes de dados, sendo ainda estabelecida uma classificação quanto à importância de cada processo e quanto à prioridade da sua implementação. Os dirigentes

intermédios são considerados os especialistas na sua área de intervenção, sendo de importância fundamental a sua validação na definição da versão final da matriz.

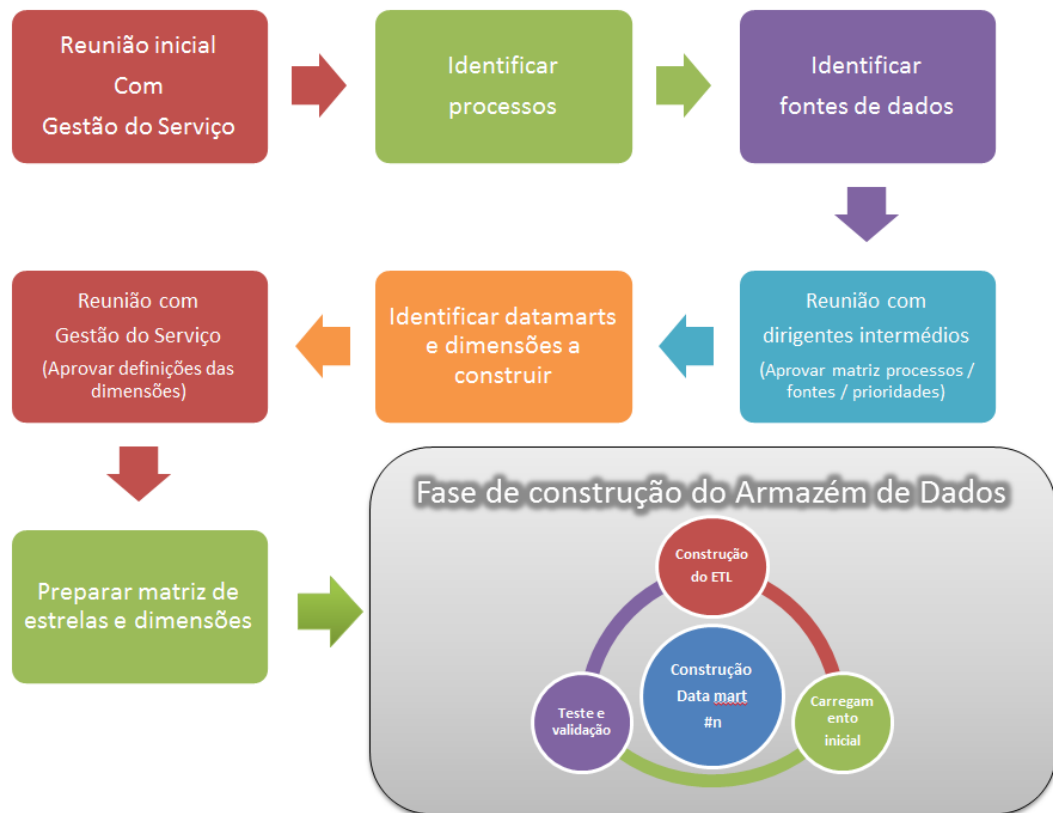


Figura 3.1.1: Metodologia de implementação do armazém de dados

Tendo como ponto de partida a matriz de processos e fontes, iniciar-se-á a fase de identificação do conjunto de *data marts* e dimensões do armazém de dados, validada através de nova reunião com a direção de serviços. Nesta fase será também importante obter um modelo de análise do serviço que apresente um conjunto de indicadores e gráficos pretendidos.

Com base na informação compilada, proceder-se-á por fim à preparação de uma matriz de estrelas e dimensões, que estará na base da definição e proposta de um modelo de dados dimensional global de um Serviço de Documentação e Informação de uma Instituição de Ensino Superior.

A fase de implementação do armazém de dados, como já aqui foi referido, segue uma metodologia iterativa cíclica, composta por três fases distintas, repetidas para cada *data mart* construído.

O processo inicia-se com a modelação de um *data mart*, selecionado pela importância e pela prioridade definidas anteriormente, a partir da matriz de barramento dimensional pré-estabelecida. Com base na informação relativa às origens de dados, proceder-se-á ao desenvolvimento dos processos de extração, transformação e carregamento de dados. Após a conclusão deste desenvolvimento, proceder-se-á ao carregamento inicial do *data mart* e ao teste e validação dos dados previamente carregados. Se os dados carregados no armazém de dados satisfizerem os critérios de qualidade e consistência pretendidos, então passar-se-á ao *data mart* seguinte, caso contrário, o processo será repetido até que todos os critérios de qualidade de dados sejam satisfeitos.

Este processo cíclico repetir-se-á até que todos os *data marts* estejam implementados ou até que o horizonte temporal da dissertação seja atingido.

4 Fase de Levantamento de Requisitos

Este capítulo dá conta do trabalho de análise realizado à estrutura organizativa do Serviço de Documentação e Informação da Faculdade de Engenharia da Universidade do Porto, marcando desta forma o início da fase de levantamento de requisitos com vista à implementação do armazém de dados.

Na primeira secção pode ser encontrada uma descrição pormenorizada do SDI, bem como de toda a sua estrutura organizativa. Na segunda secção é apresentado o levantamento de processos efetuados e respetivos casos de uso. Na terceira secção são elencadas as fontes de dados disponíveis para a construção do armazém de dados. Por fim, na quarta secção, é apresentada a matriz que cruza os processos com as fontes de dados e que no fundo sumariza todo o trabalho de investigação elaborado nesta fase e permitirá passar à fase seguinte, onde serão identificados os *data marts* e dimensões do futuro armazém de dados.

4.1 Análise Orgânico Funcional do SDI

O Serviço de Documentação e Informação da FEUP (SDI) estava até 2009 estruturado organicamente seguindo um modelo organizativo hierárquico, assente na tipologia dos recursos informacionais existentes: Biblioteca, Arquivo e Museu.

A regulamentação orgânica do Serviço de Documentação e Informação da Faculdade de Engenharia da Universidade do Porto (FEUP) foi publicada em Diário da República a 9 de Agosto de 2004. Aí se determina a sua área de atuação especificando que os SDI “exercem a sua atividade no âmbito da conceção, da gestão, do tratamento, da difusão e da conservação de documentação e informação científico-técnica e de cariz pedagógico, visando o apoio ao ensino e à investigação e a preservação da memória e do património cultural e tecnológico da FEUP” (Faculdade de Engenharia da Universidade do Porto 2004, 3).

Também no mesmo regulamento se especifica a sua estrutura orgânica, sendo os SDI constituídos por três Divisões e uma Unidade.

As três Divisões são:

- Divisão de Arquivo e Museu
Esta Divisão tem a sua área de atuação no “âmbito da gestão da informação administrativa e do património documental e museológico” (Faculdade de Engenharia da Universidade do Porto 2004, 3)

- Divisão de Biblioteca
A Divisão de Biblioteca tem a sua área de atuação circunscrita ao “âmbito da seleção, do tratamento e do acesso à informação documental de cariz científico-técnico e pedagógico” (Faculdade de Engenharia da Universidade do Porto 2004, 4)
- Divisão de Serviços Eletrónicos
Exerce a sua atividade no âmbito “da seleção, do tratamento e do acesso à informação documental de cariz científico-técnico e pedagógico.” (Faculdade de Engenharia da Universidade do Porto 2004, 4)

A Unidade é:

- FEUP Edições
Exerce a sua atividade “no âmbito da gestão da editora da FEUP” tendo como finalidade “Difundir as atividades de I&D produzidas pela comunidade docente, investigadora e colaboradora da Faculdade” e “Produzir materiais pedagógicos inovadores de suporte ao ensino e à aprendizagem” (Faculdade de Engenharia da Universidade do Porto 2004, 4)

Todas as Divisões e a Unidade dispõem de “missões, recursos e pessoal específico” (Faculdade de Engenharia da Universidade do Porto).

Em 2009 ocorre uma reestruturação do Serviço que, não alterando a sua estrutura orgânica no plano formal, provoca uma rutura com o modelo hierárquico vigente, introduzindo um novo modelo de organização matricial, introduzindo alterações no plano da organização do trabalho, seguindo uma lógica de organização orientada para os serviços de informação prestados aos utilizadores (AGORA 2009; ESA 2009).

São assim criadas duas áreas transversais ao Serviço, a Área de Serviços Diretos à Comunidade e a Área de Gestão Documental (AGORA 2009).

São ainda constituídas 6 equipas também elas transversais às divisões formalmente existentes, inseridas nas áreas acima mencionadas de forma a concretizar as diferentes atividades daí decorrentes.

A Figura 4.1.1 representa a estrutura orgânica formal do SDI sobreposta pela organização de trabalho transversal emanada da reestruturação de 2009.

Apresenta-se a seguir uma breve descrição das seis equipas de trabalho existentes. Posteriormente serão apresentados os casos de uso que tentam caracterizar as atividades e os processos nos quais essas equipas estão envolvidas e que no fundo são responsáveis por dados registados em sistemas e que poderão ser analisados com a finalidade de serem aproveitados para a construção do armazém de dados. Este trabalho é portanto realizado no pressuposto de uma gestão por processos do SDI.

AGORA

A equipa AGORA organicamente enquadra-se “na Área de Serviços Diretos à Comunidade, estando particularmente vocacionada para o desempenho de atividades de *front office* que concretizem a missão dos SDI no que toca à exploração e à difusão do seu património científico e cultural junto da comunidade interna e externa à FEUP, e que promovam a imagem dos SDI como conjunto de serviços de apoio às atividades pedagógicas, de investigação, e administrativas da mesma instituição” (AGORA 2009, 3).

A equipa AGORA tem como missão “inspirar e promover formas inovadoras de olhar, pensar e agir sobre o património científico-tecnológico e a cultura, a partir da criação de momentos-fórum nos quais o questionamento, o contacto e as novas abordagens de matérias atuais do foro do conhecimento e da cidadania se possam desenvolver com abertura e criatividade” (AGORA 2009, 3).

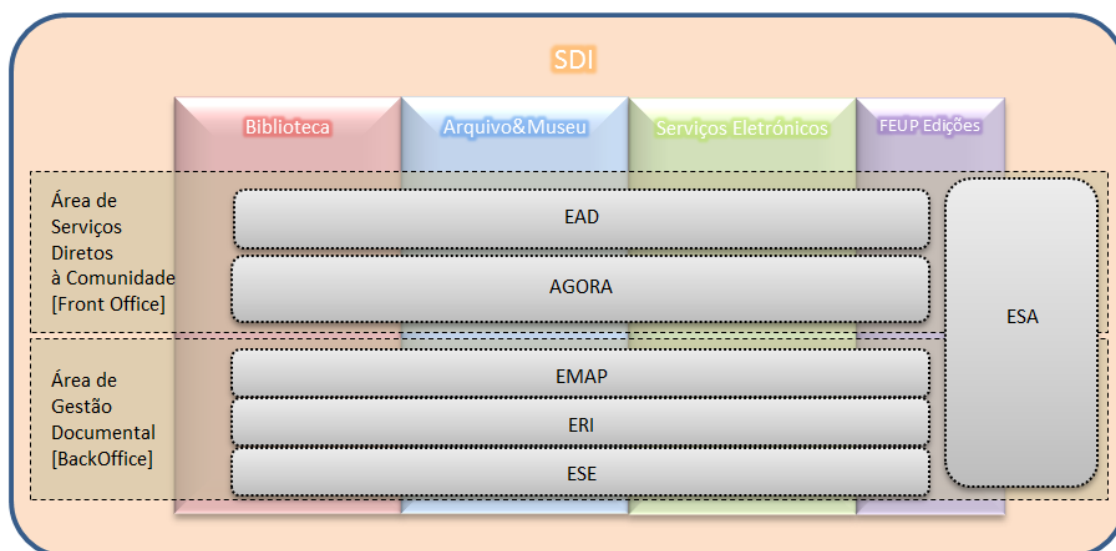


Figura 4.1.1: Representação da estrutura orgânica formal do SDI sobreposta pela organização de trabalho transversal emanada da reestruturação de 2009.

EQUIPA DE APOIO E DESCOBERTA (EAD)

Segundo o seu documento identitário, a EAD “compromete-se a ligar a comunidade da FEUP com a informação de que ela precisa, numa preocupação constante em a sensibilizar e dotar com as competências de infoliteracia que são necessárias para a autonomização da sua aprendizagem, mesmo para além da sua vida académica. Para o conseguir, a EAD assegura serviços e produtos de informação que são desenvolvidos em função da comunidade da FEUP, com a qual pretende manter uma estreita e contínua cooperação” (EAD 2009, 5).

EQUIPA DE MEMÓRIA E ACESSO PERENES (EMAP)

A EMAP tem como missão “Desenvolver uma política integrada de conservação e preservação, difusa a toda a informação produzida na FEUP, de âmbito pedagógico, científico, administrativo e patrimonial, sem discriminação de suporte ou ambiente tecnológico de base, de modo a garantir no tempo a persistência do seu acesso e recuperação, salvaguardando, assim, os respetivos valores probatórios, legais, informativos e a memória dos atos e história da Faculdade” (EMAP 2009, 4).

EQUIPA DE REPRESENTAÇÃO DE INFORMAÇÃO (ERI)

Esta equipa “é a equipa responsável pela representação eficiente da informação contida e registada nos diferentes suportes materiais à guarda do SDI, nomeadamente em documentos de diferentes tipologias como Biblioteca, Arquivo e Museu. A sua missão assenta no conhecimento técnico da representação rigorosa desta informação, integrando as diversas formas do seu registo e uniformizando as metodologias do seu tratamento, com vista à sua eficaz recuperação e disponibilização” (ERI 2009, 3).

EQUIPA DE SUPORTE À ADMINISTRAÇÃO (ESA)

Organicamente “a ESA integra-se na área de suporte infraestrutural de serviços e de recursos, e nas áreas de administração e gestão. A ESA tem estruturas funcionais bem definidas, com referência a áreas interfuncionais e interequipas, que denotam a sua natureza de direção, de suporte e de avaliação” (ESA 2009, 3).

A equipa tem por missão “garantir os requisitos necessários ao cumprimento da Missão do SDI, recorrendo a práticas adequadas de gestão, para integrar e otimizar o Capital Natural – envolvimento, contexto e ambiente, o Capital Físico – infraestruturas, equipamentos, sistema financeiro, o Capital Humano – nível de competências e base de conhecimento, e o Capital Social – coesão, confiança, participação, existentes neste Serviço” (ESA 2009, 3).

EQUIPA DE SERVIÇOS ELETRÓNICOS (ESE)

A missão da ESE consiste em exercer “a sua atividade no âmbito da produção e da gestão das coleções digitais, dos serviços web e dos sistemas informáticos específicos da direção de serviços” (ESE 2009, 7).

4.2 Processos e Casos de Uso

Com base na análise feita aos documentos identitários foi possível identificar os principais processos nos quais as equipas estão envolvidas, as principais atividades que consubstanciam cada processo e ainda onde o desenvolvimento dessas atividades potencialmente regista dados que possam ser aproveitados para a construção do armazém de dados.

Também com base nessa informação, foram elaborados os casos de uso de cada processo, procurando evidenciar assim as interações existentes com outras equipas e com outras entidades externas ao Serviço.

Assim, para cada equipa, foi construída uma tabela semelhante à Tabela 1, com a listagem dos seus processos, as suas atividades e ainda uma breve descrição das mesmas, seguindo-se a representação dos respetivos diagramas de casos de uso, semelhante à Figura 4.1.1, Figura 4.2.2 e Figura 4.2.3 apresentadas em baixo. O caso específico, ao qual o exemplo se reporta, representa a lista de processos e casos de uso da equipa AGORA. A informação deste teor referente às outras equipas poderá ser consultada no Anexo C.

Tabela 1: Sumário dos processos/atividades da equipa AGORA

Processos	Atividades	Descrição
Coordenação	• Conceber projeto	Conceção de eventos de cariz cultural.
	• Gerir projeto	Gerir todas as fases da construção de um projeto.
Interpretação de documento	• I&D&I	Investigação em torno de um documento.
	• Recolher informação	Recolha de informação relacionada com um documento.
	• Conceber proposta de interpretação	Conceção de um documento que resume o trabalho de I&D&I desenvolvido em torno de determinado documento.
Eventos	• Programar	Elaboração do programa do evento.
	• Produzir	Produção do evento, seguindo a programação previamente estipulada.

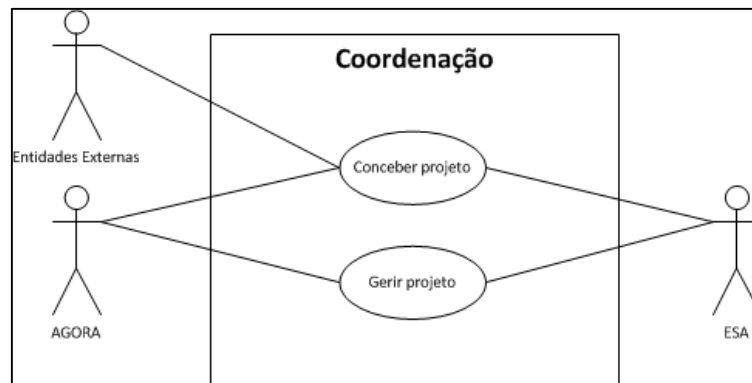


Figura 4.2.1 Casos de Uso do processo Coordenação

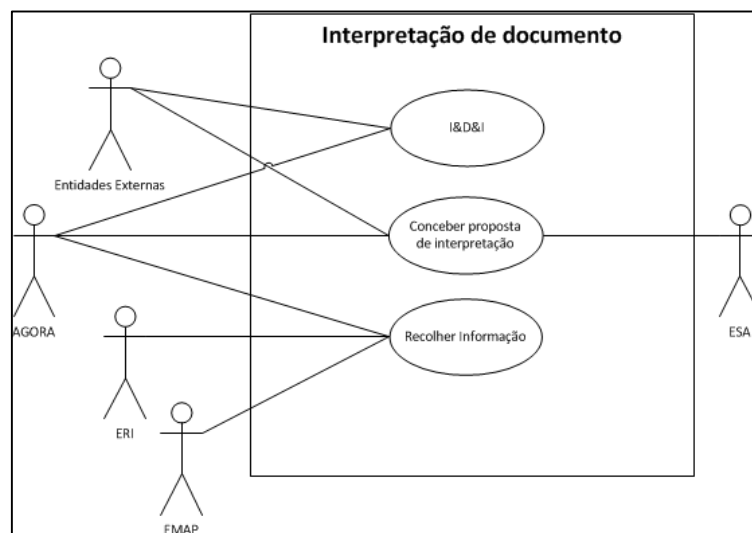


Figura 4.2.2. Casos de Uso do processo Interpretação de Documento

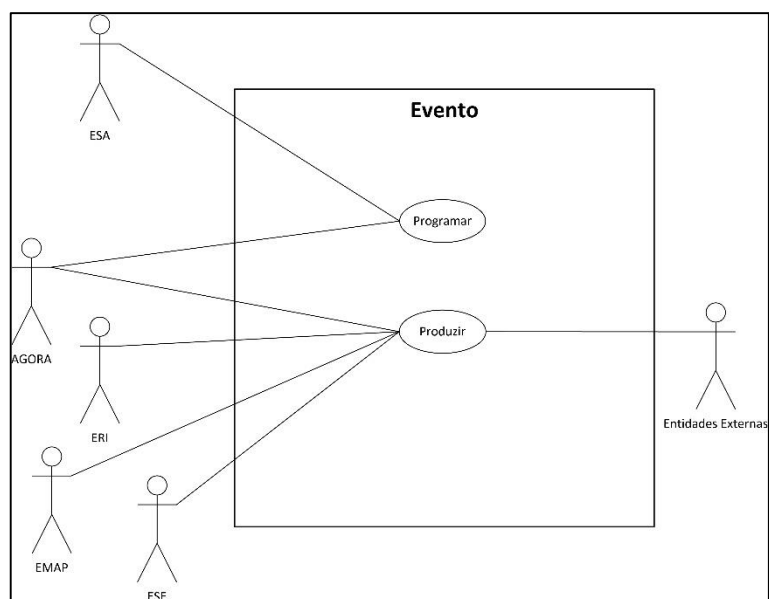


Figura 4.2.3 Casos de Uso do processo Evento

4.3 Fontes de dados

Apresenta-se em baixo, sob a forma de quadro resumo, o levantamento efetuado às fontes de dados que de alguma forma estão envolvidas nos processos identificados na análise anterior.

A identificação das fontes de dados, foi conseguida não só devido ao trabalho de análise efetuada, mas também ao conhecimento do autor nesta matéria, graças à experiência profissional acumulada como administrador dos sistemas informáticos do SDI. Paralelamente, conversas informais com as chefias intermédias do SDI ajudaram a complementar a lista que a seguir se apresenta.

Tabela 2: Tabela de fontes de dados

Fonte de dados	Descrição
Aleph	Sistema de Gestão Integrada de Bibliotecas. Sistema operacional que faz a gestão da informação da biblioteca na sua faceta mais tradicional, cobrindo as áreas da catalogação e inventário, circulação e empréstimo de material, aquisições e controlo de periódicos, empréstimo interbibliotecas e catálogo em linha de acesso público.
Digitool	O Digitool é o repositório institucional da FEUP, onde são armazenados todos os conteúdos maduros, fruto da produção académica, científica e de investigação. Isto inclui dissertações de mestrado e teses de doutoramento, trabalhos de alunos, publicações dos docentes e investigadores, documentos do arquivo, objetos digitais do museu, fotografias institucionais e de eventos, etc...
Metalib	Portal de meta-pesquisa da FEUP. Permite numa só caixa de entrada lançar pesquisas a vários recursos de informação em simultâneo.
SFX	O SFX é um resolvidor de ligações. Permite construir ligações que dão acesso ao texto integral de recursos de informação. É um servidor de ligações contextuais para informação relacionada, após efetuar uma análise automática aos dados de uma pesquisa.
InArte	Sistema de Gestão Integrada de Museus. Sistema operacional que faz a gestão de informação do Museu da FEUP. Conta com as funcionalidades de gestão da inventariação, conservação e circulação de artefactos. Conta também com um catálogo de acesso em linha.
GISA	Sistema de Gestão Integrada de Arquivos. Sistema operacional que faz a gestão de informação do Arquivo da FEUP. Conta com as funcionalidades de gestão da inventariação, conservação e circulação de documentação. Conta também com um catálogo de acesso em linha.
Recursos Eletrónicos	Estão incluídos nesta categoria todos os fornecedores de recursos de informação contratados pela FEUP. Nela incluem-se editores de revistas, e-books, normas, patentes, agregadores de informação (com e sem acesso a texto integral), etc...
Moodle	O Moodle é a plataforma de ensino à distância adotada na FEUP.
Sharepoint	O Sharepoint é a solução adotada na FEUP para suportar as intranets da Biblioteca e do CICA.
SIGARRA	O SIGARRA é o Sistema de Informação para Gestão Agregada dos Recursos e dos Registos Académicos.
Seagull	O Seagull é um gestor de conteúdos, usado para a construção de sítios WEB. No SDI está na base do seu atual portal WEB.

Fonte de dados	Descrição
Drupal	O <i>Drupal</i> é um gestor de conteúdos, usado para a construção de sítios WEB. No SDI estará na base do seu futuro portal WEB e suporta já vários portais temáticos.
MediaWiki	Plataforma colaborativa que alberga as páginas do guia de apoio à publicação da FEUP.
Open Journal System	O OJS é uma aplicação baseada em código aberto que faz a gestão da publicação de revistas em linha.
GRH	A aplicação de Gestão de Recursos Humanos usada na FEUP.
QualiFEUP	Manuais elaborados de forma colaborativa no <i>contribute</i> .
Site FEUP Edições	Sítio <i>web</i> da Editora da FEUP.
Redmine	A Redmine é um gestor de projetos <i>web</i> baseado em código aberto.
Email	A infraestrutura de processamento e armazenamento de mensagens de correio eletrónico adotada na FEUP.
Google Books	É um serviço do Google que pesquisa no texto integral de livros.
Outras fontes	Conjunto de fontes de dados ad hoc, onde se incluem ficheiros de texto, folhas de cálculo, etc...

4.4 Matriz de Processos e Fontes de Dados

Apresenta-se a seguir a matriz que no fundo faz o resumo da investigação efetuada e descrita nas três secções anteriores. Dada a extensão da matriz, a mesma será partida em unidades mais pequenas, uma matriz para cada equipa do SDI, de forma a melhorar a sua leitura. O exemplo apresentado na Tabela 3, em baixo, reporta-se à matriz de processos e fontes de dados da equipa ERI. As restantes matrizes poderão ser consultadas no Anexo D

EQUIPA DE REPRESENTAÇÃO DE INFORMAÇÃO (ERI)

Tabela 3: Matriz processos/fontes de dados para as atividades da equipa ERI

Processos	<div>Atividades</div> <div>Fontes</div>		Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Fontes ad hoc
Gestão documental	Tratamento documental		●	●			●	●			●												
	Disponibilização de Documentos		●	●			●	●													●		
Tarefas de Gestão	Planeamento e gestão de tarefas, reuniões,										●	●									●		●
	Controlo de qualidade											●									●		
	Gerir a base de autoridade		●				●	●				●						●					
	Gerir as assinaturas correntes		●																				

4.5 Conclusões

Neste capítulo apresentaram-se os trabalhos levados a cabo e que marcaram o arranque da fase de levantamento de requisitos com vista ao desenvolvimento do armazém de dados.

Foi apresentado um estudo orgânico-funcional do SDI, com o objetivo de ficar a conhecer pormenorizadamente a orgânica e funcionamento do serviço e fazer um levantamento exaustivo de todos os seus processos, atividades e atores responsáveis pela criação de dados.

Com esta abordagem tenta-se garantir que nenhum facto importante fica à partida excluído do processo de desenho e conceção deste armazém de dados.

Foram identificados 21 processos e 93 casos de uso que foram devidamente representados sob a forma de diagramas UML.

Posteriormente foram elencadas as fontes de dados disponíveis para servirem como origem de dados de um futuro armazém de dados. A lista apresentada contou com as contribuições do autor deste documento (dada a sua experiência profissional acumulada como administrador dos sistemas do SDI), da análise efetuada e ainda de conversas informais levadas a cabo com as chefias intermédias do SDI.

Por fim foi apresentada a matriz que faz o cruzamento entre os processos e as fontes de informação disponíveis.

Estas atividades estão em conformidade com a metodologia expressa para a prossecução deste trabalho de investigação, abrindo caminho para a próxima fase, também ela expressa na metodologia, que se resume à identificação dos *data marts* e das dimensões que irão fazer parte do armazém de dados.

5 Projeto

Na fase de levantamento de requisitos foi levado a cabo um estudo sobre a estrutura orgânica do Serviço de Documentação e Informação, que incluiu uma análise aos processos, casos de uso e fontes de dados envolvidos nas atividades do Serviço. Esta fase culminou na construção de uma matriz que resume a arquitetura de processos do Serviço, na perspetiva das fontes de dados que os suportam. A referida matriz foi posteriormente discutida e validada em reuniões realizadas com a Direção de Serviços e com os Coordenadores das Equipas.

Com base no trabalho desenvolvido no capítulo anterior, definem-se agora os indicadores relevantes, importâncias e prioridades de implementação, culminando o capítulo com a apresentação do projeto de um modelo dimensional a implementar.

5.1 Levantamento de indicadores

Com base na matriz de processos e fontes de dados construída, iniciou-se uma fase de reuniões individuais com as chefias intermédias com o objetivo de recolher junto destas um conjunto de indicadores e gráficos que seria relevante obter após a implementação do armazém de dados. A Tabela 4, em baixo apresentada, sintetiza o trabalho de levantamento dos indicadores, distribuídos pelas equipas e pelos processos que lhes estão afetos. Foram recolhidos um total de 127 indicadores, distribuídos por 21 processos e 6 equipas. Aos indicadores foi atribuída uma numeração estruturada para a sua melhor identificação.

Tabela 4: Lista de indicadores, distribuídos por processo e equipa

Equipa	Processo	Indicadores
1. AGORA	1.1. Coordenação	1.1.1. Tempo gasto na conceção de projeto;
		1.1.2. Tempo gasto no desenvolvimento das atividades;
	1.2. Interpretação	1.2.1. Tempo gasto no trabalho de interpretação;
	1.3. Eventos	1.3.1. Relações estabelecidas entre SDI/Departamentos e Serviços FEUP (ou seja, mapear as relações de forma a assegurar a representatividade de todos os Departamentos/Serviços nas atividades culturais);
		1.3.2. Índices de participação nos eventos;

Equipa	Processo	Indicadores
		1.3.3. Satisfação do participante; 1.3.4. Impacto das atividades sobre movimento dos acervos;
2. EAD	2.1. FrontOffice	2.1.1. Nº de utilizadores inscritos; 2.1.2. Nº de novos utilizadores inscritos; 2.1.3. Nº de documentos requisitados; 2.1.4. Nº de documentos devolvidos; 2.1.5. Nº de documentos renovados; 2.1.6. Nº de documentos reservados; 2.1.7. Multas cobradas; 2.1.8. Utilizadores que fizeram empréstimos; 2.1.9. Informação sobre empréstimos em atraso (n.º de obras extraviadas); 2.1.10. Número de empréstimos por piso (por posto de trabalho e por obra); 2.1.11. Requisições ao arquivo, totais, n.º requisições, n.º de documentos; 2.1.12. Arrumação dos documentos devolvidos ao Arquivo (N.º de documentos devolvidos); 2.1.13. Nº de utilizadores atendidos; 2.1.14. Nº de questões colocadas e respondidas (Interessa segmentar por tipo ou meio de apoio (presencial, e-mail, telefone)); 2.1.15. Nº de consultas ao arquivo, total, de administração, de investigação, manutenção de arquivos correntes; 2.1.16. Nº de pedidos de informação e pesquisa ao Arquivo, por tipo de contato, serviço e operador; 2.1.17. Nº de obras consultadas (consulta presencial); 2.1.18. Uso dos computadores; 2.1.19. Uso dos lugares de leitura; 2.1.20. Ocupação dos pisos; 2.1.21. Uso dos cacifos; 2.1.22. Uso dos gabinetes; 2.1.23. Pedidos de reserva de gabinetes aprovados; 2.1.24. Pedidos de reserva de gabinetes recusados; 2.1.25. Uso das fotocopiadoras;
	2.2. Infoliteracia	2.2.1. Nº de sessões de formação realizadas; 2.2.2. Nº de formandos que participaram nas sessões; 2.2.3. Nº de horas de formação; 2.2.4. Indicadores iguais aos de 1ª linha. (Interessa segmentar por tipo de utilizadores e por meio de apoio (presencial, e-mail, telefone)); 2.2.5. Nº de utilizadores externos inscritos; 2.2.6. Uso dos recursos. Sessões, pesquisas e visualizações, por recurso e categoria (bases de dados bibliográficas, e-books, aplicações...);
	2.3. Tarefas de gestão	2.3.1. Nº de trouble tickets colocados; 2.3.2. Nº de trouble tickets respondidos; 2.3.3. Nº de entradas biblioteca; 2.3.4. Nº de documentos rececionados;

Equipa	Processo	Indicadores
		2.3.5 N° de documentos não localizados; 2.3.6 N° de documentos danificados; 2.3.7 N° de problemas comportamentais registados; 2.3.8 N° de utilizadores sancionados;
3. EMAP	3.1. Tratamento Técnico de Preservação	3.1.1. Monitorização ambiental (humidade relativa, temperatura e luminosidade); 3.1.2. Higienização de documentos (número de documentos limpos) (ofertas ocasionais e atividade regular sobre os acervos dos depósitos); 3.1.3. Objetos museológicos e documentos intervencionados quanto a atos de conservação/acondicionamento; 3.1.4. Manutenção de arquivos correntes no Arquivo (n.º de processos intervencionados); 3.1.5. Transferências, incorporações e alienações no Museu; 3.1.6. Transferências, incorporações e alienações no Arquivo; 3.1.7. Encadernação de documentos;
	3.2. Gestão Documental Física	3.2.1. Transferências, incorporações e alienações na Biblioteca; 3.2.2. Ofertas de objetos museológicos, por tempo, por entidade ofertante; 3.2.3. Taxa de ocupação de depósitos, por entidade produtora, crescimento anual. A medida é o metro linear; 3.2.4. Instalação/reinstalação de documentos e objetos museológicos, em depósito e reserva;
4. ERI	4.1. Gestão Documental	4.1.1. Registos bibliográficos criados na Biblioteca; 4.1.2. Registos bibliográficos alterados na Biblioteca; 4.1.3. Registos de exemplares criados na Biblioteca; 4.1.4. Registos de exemplares alterados na Biblioteca; 4.1.5. Registos de autoridade criados na Biblioteca; 4.1.6. Registos de autoridade alterados na Biblioteca; 4.1.7. Duração das diversas fases do tratamento técnico; 4.1.8. Receção de documentação para tratamento técnico; 4.1.9. Registo de documentos novos no Arquivo (unidades informacionais - GISA); 4.1.10. Alterações a documentos no Arquivo (unidades informacionais - GISA); 4.1.11. Registos de exemplares criados no Arquivo (unidades físicas - GISA); 4.1.12. Alterações a exemplares no Arquivo (unidades físicas - GISA); 4.1.13. Registos de autoridade criados no Arquivo (GISA); 4.1.14. Registos de autoridade alterados no Arquivo (GISA);

Equipa	Processo	Indicadores
		4.1.15.Registo de documentos novos no Museu (INARTE); 4.1.16.Alterações a documentos no Museu (INARTE); 4.1.17.Registo de exemplares criados no Museu (INARTE); 4.1.18.Alterações a exemplares no Museu (INARTE); 4.1.19.Registos de autoridade criados no Museu (INARTE); 4.1.20.Registos de autoridade alterados no Museu (INARTE);
	4.2. Tarefas de Gestão	4.2.1. Controlo de qualidade. Número de não conformidades encontradas no decorrer de inspeções efetuadas às descrições dos documentos; 4.2.2. Indicador relativo à participação da ERI em outras tarefas/projetos/eventos;
5. ESA	5.1. Coordenação	5.1.1. Percentagem de execução do plano de atividades; 5.1.2. Número de tarefas do plano de atividades concluídas; 5.1.3. Visitas oficiais à Biblioteca; 5.1.4. Instruções de trabalho novas; 5.1.5. Instruções de trabalho corrigidas; 5.1.6. Projetos de cooperação e integração a nível nacional e internacional; 5.1.7. Indicadores sobre reclamações dos utentes;
	5.2. Secretariado Técnico	5.2.1. Número de mensagens de correio eletrónico de divulgação enviadas; 5.2.2. Número de publicações no <i>website</i> e página facebook do serviço; 5.2.3. Número de notícias publicadas no SIGARRA; 5.2.4. Número de iniciativas publicitárias colocadas na área dos elevadores;
	5.3. Gestão de Infraestruturas	5.3.1. Condições ambientais da Biblioteca; 5.3.2. Gestão de espaços [verificação do estado de conservação e funcionamento dos espaços e equipamentos - do número de ocorrências]; 5.3.3. Gestão de inventário [número de cadeiras, mesas, computadores. Bem confirmados/Não confirmados....];
	5.4. Gestão de Pessoas	5.4.1. Assiduidade; 5.4.2. Reuniões – periodicidade;
	5.5. Gestão Financeira	5.5.1. Volume de custos efetuados por diversas rubricas previamente definidas, por ex. bibliografia, material escritório, formação, e-books, etc...; 5.5.2. Informação semanal sobre o total gasto, recebido e saldo do SDI entregue à direção em reunião semanal;
	5.6. Gestão de Aquisições	5.6.1. Bibliografia das unidades curriculares que existem na biblioteca ou não;

Equipa	Processo	Indicadores
		5.6.2. EIB - Tempo de fornecimento de documentos (número de pedidos, tempos mínimos, médios e máximos por tipo de material); 5.6.3. Tempo de aquisição de documentos; 5.6.4. Aquisição ao preço mais baixo. Listagem dos fornecedores e número de aquisições efetuadas e valor transacionado, por tipo de produto; 5.6.5. N.º de obras extraviadas repostas;
	5.7. Gestão FEUP Edições	5.7.1. Número de exemplares vendidos, por Título; 5.7.2. Valor das vendas, global, por título, por cliente; 5.7.3. Tempo de preparação de obras; 5.7.4. Tempo de produção de obras;
6. ESE	6.1. Produção de documentos	6.1.1. N.º de documentos eletrónicos produzidos; 6.1.2. N.º de documentos depositados no Digitool; 6.1.3. Tempo de disponibilização dos documentos; 6.1.4. Tempo de produção de <i>e-books</i> ; 6.1.5. Satisfação do cliente e do responsável da coleção quanto ao <i>e-book</i> produzido;
	6.2. Projetos	6.2.1. Tempo gasto no desenvolvimento de funcionalidades;
	6.3. Gestão de Serviços WEB	6.3.1. Satisfação do cliente quanto à produção de sítios temáticos; 6.3.2. Tempo de desenvolvimento dos serviços e materiais de comunicação; 6.3.3. Dados analíticos do comportamento dos utilizadores – Percurso; 6.3.4. Dados analíticos do comportamento dos utilizadores - Ponto de acesso; 6.3.5. Dados analíticos do comportamento dos utilizadores – Tempo; 6.3.6. Dados analíticos do comportamento dos utilizadores – Saída; 6.3.7. Dados analíticos do comportamento dos utilizadores - Páginas mais usadas;
	6.4. Gestão da infraestrutura tecnológica	6.4.1. Taxa de disponibilidade dos serviços; 6.4.2. Número de reclamações; 6.4.3. Tempo gasto no desenvolvimento de funcionalidades nas aplicações de suporte; 6.4.4. Tempo gasto na manutenção da infraestrutura; 6.4.5. Satisfação do cliente quanto às soluções desenvolvidas; 6.4.6. Indicador de ocupação de espaço em disco; 6.4.7. Número de problemas colocados; 6.4.8. Tempo de resposta aos TTs relacionados com a infraestrutura tecnológica;
TOTAIS: EQUIPAS: 6	PROCESSOS: 21 INDICADORES LEVANTADOS: 127	

5.2 Importância e prioridade de implementação

Paralelamente ao trabalho de recolha de indicadores tentou-se ainda determinar, do ponto de vista dos coordenadores, qual seria a importância atribuída a cada processo, tendo em conta o universo da sua equipa. Para tal foi-lhes pedido que classificassem a importância de cada um deles, recorrendo a uma escala numérica de valores discretos, compreendidos entre 1 (mais importante) e 5 (menos importante).

Dada a complexidade para a implementação total de um armazém de dados com estas características foi ainda criado um segundo nível de classificação no qual se pretende clarificar quais os processos mais prioritários a implementar. Nesta classificação foi usada uma escala numérica de valores discretos, compreendidos entre 1 (mais prioritário) e 3 (menos prioritário).

Uma reunião geral com chefias intermédias e a direção de serviços serviu para validar e harmonizar toda a informação recolhida.

Tendo em conta os critérios em cima especificados e ainda a premissa de que o esforço de desenvolvimento deveria suportar pelo menos a implementação de um *data mart* para cada equipa, seleccionaram-se os processos assinalados a fundo verde na Tabela 5 abaixo apresentada, perfazendo um total de 8 processos em 21 possíveis.

Tabela 5: Lista de processos acompanhada da sua importância e prioridade de implementação

Equipa	Processo	Importância [1..5]	Prioridade [1..3]
AGORA	Coordenação	2	2
	Interpretação	3	3
	Eventos	1	1
EAD	FrontOffice	1	1
	Infoliteracia	1	1
	Tarefas de gestão	3	2
EMAP	Gestão Documental Física	3	2
	Tratamento Técnico de Preservação	3	3
ERI	Gestão Documental	1	1
	Tarefas de Gestão	5	3
ESA	Coordenação	3	2
	Secretariado Técnico	4	3
	Gestão de Infraestruturas	2	2
	Gestão de Pessoas	4	3
	Gestão Financeira	1	1
	Gestão de Aquisições	1	1
	Gestão da FEUP Edições	5	3
ESE	Produção de documentos	2	1
	Projetos	4	3
	Gestão de Serviços WEB	3	3
	Gestão da infraestrutura tecnológica	2	2
TOTAIS:	EQUIPAS: 6	PROCESSOS LEVANTADOS: 21	PROCESSOS A IMPLEMENTAR: 8

5.3 Modelo dimensional

Com base na informação compilada em seções anteriores, procedeu-se à construção do modelo dimensional do armazém de dados, baseado nas metodologias especificadas por Kimball (1998; 2002).

Este processo inicia-se com a construção da matriz de barramento dimensional, onde todos os *data marts* e dimensões de possível implementação são determinados e discriminados. Neste estudo de caso específico, cada *data mart* corresponde a um processo de negócio identificado na fase de levantamento de requisitos, podendo também ser visto como uma coleção de estrelas que pretende dar resposta aos indicadores identificados na seção 5.1 acima.

A matriz de barramento dimensional resultante deste processo de análise encontra-se disponível para consulta no Anexo E, dada a sua considerável extensão. Esta matriz funciona como um quadro de referência para o desenvolvimento de cada *data mart*.

Dada a natureza iterativa do processo de conceção do modelo dimensional, a matriz apresentada não é de todo um modelo fechado, existindo uma forte probabilidade de algumas dimensões inicialmente apresentadas virem a tornar-se meras dimensões degeneradas (dimensões sem qualquer tipo de atributo) ou até mesmo a desaparecerem por se chegar à conclusão que não são necessárias. O mesmo poderá suceder à lista de estrelas apresentadas na matriz. É importante que a definição de cada dimensão permaneça consolidada no barramento para evitar incompatibilidades entre as várias estrelas. Como estas definições representam conceitos importantes na organização, determinantes na respetiva avaliação, devem ser explicitamente aprovados pela gestão. A dimensão “documento” é paradigmática, tendo o seu conceito sido discutido no decorrer dos trabalhos desta dissertação. Um documento pode ser visto de forma diversa, dependendo da área a partir do qual ele é abordado. Ou seja, um documento para uma biblioteca assume um significado que não será necessariamente coincidente com a perspetiva de um museu ou de um arquivo. No caso concreto do trabalho realizado nesta dissertação, a discussão em torno do conceito de “documento” levou a que a dimensão fosse implementada com um atributo que permitisse distinguir a natureza do mesmo e assim obter indicadores úteis para cada unidade, bem como para todo o serviço.

Uma vez identificados todos os possíveis *data marts* e suas dimensões associadas, foi aplicado o método dos quatro passos, também preconizado por Kimball (1998; 2002) para a especificação das tabelas de factos.

O método dos quatro passos inicia-se com a escolha de um *data mart* e dentro desse *data mart*, a escolha da tabela de factos candidata a ser especificada, passo 1.

No passo 2 é determinada a granularidade dessa tabela de factos. Este passo reveste-se de crucial importância, uma vez que o sucesso dos dois últimos passos está dependente da precisão com que a granularidade da tabela de factos é determinada. Apurar a granularidade de uma tabela de factos significa especificar exatamente o que uma linha de uma tabela de factos representa.

No passo 3 são escolhidas as dimensões que melhor se adequam à tabela de factos, tendo em conta a granularidade especificada no passo 2. A par com a determinação das dimensões, a lista dos seus atributos é também especificada.

No quarto e último passo são determinados os factos propriamente ditos, as medidas representadas na tabela de factos. Todos os factos devem fazer sentido para a granularidade especificada no passo 2. Factos de granularidade diferente pertencem a tabelas de factos diferentes. Tipicamente os factos são numéricos e aditivos, embora seja possível a existência de factos não numéricos.

A semelhança da matriz de barramento, a lista das tabelas de factos apuradas é uma lista extensa, podendo ser consultada na sua totalidade no Anexo F. Em baixo, na Tabela 6, é apresentado um exemplo da especificação de uma tabela de factos, onde é possível observar pormenorizadamente todos os seus atributos: a granularidade da tabela de factos, as dimensões associadas e respetiva descrição e por fim as medidas que essa tabela de factos pretende capturar. As dimensões degeneradas são devidamente identificadas através do código “(DD)”.

Tabela 6: Exemplo de especificação de uma tabela de factos

Estrela	Multas	Versão	2.1	Data	25/05/2013
Granularidade	Multa cobrada a utilizador, resultante da sua atividade... multas, inscrições, etc...				
Dimensões					
Documento	Documento				
Colaborador	Colaborador que cobra a multa				
Utilizador	A quem é cobrada a multa				
Estado (DD)	Estado da transação (cobrada, não cobrada)				
Tipo (DD)	Tipo de transação (multas, inscrições)				
Recibo (DD)	Número do recibo				
Tempo	Data em que a transação foi lançada				
Tempo	Data de pagamento da multa				
Medidas					
Valor	Valor em euros de multas cobradas a utilizadores				

As dimensões são caracterizadas pela enumeração e especificação de todos os seus atributos, níveis e hierarquias, bem como o seu comportamento face à alteração de dados na dimensão (*Slowly Changing Dimension Type*). A relação de todas as dimensões identificadas pode ser encontrada no Anexo G. A título de exemplo é apresentada em baixo a Tabela 7, onde se pode observar o nível de detalhe usado na especificação de uma dimensão.

Tabela 7: Exemplo da especificação de uma dimensão

Nome	Descrição	SCD	Versão	1.0	Data	25/05/2013	Fontes
Categoria Pedido	Categorias de pedidos de apoio de 1ª linha	Tipo 1	Hierarquia	CATEGORIA>SUBCATEGORIA			
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	Sharepoint
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
CATEGORIA_ID	Chave de ligação do nível na hierarquia	CATEGORIA	LK	NUMBER			
CATEGORIA	ID da categoria de topo	CATEGORIA	BK	VARCHAR2	50		
DESCRICAO_CATEGORIA	Descrição da categoria de topo	CATEGORIA		VARCHAR2	240		
SUBCATEGORIA_ID	Chave de ligação do nível na hierarquia	SUBCATEGORIA	LK	NUMBER			
SUBCATEGORIA	ID do assunto base	SUBCATEGORIA	BK	VARCHAR2	100		
DESCRICAO_SUBCATEGORIA	Descrição do assunto base	SUBCATEGORIA		VARCHAR2	240		

Todo o trabalho desenvolvido nesta fase foi apresentado em reunião à Direção de Serviços e demais coordenadores de equipas. Durante a reunião procedeu-se à validação da matriz de barramento dimensional, dos factos, das dimensões e respetivas granularidades. O glossário de termos detetados no levantamento efetuado foi também alvo de debate (a definição encontrada no campo *Descrição*, nas tabelas constantes nos Anexo F e Anexo G).

5.4 Conclusões

Neste capítulo apresentaram-se os trabalhos realizados na fase de projeto do armazém de dados. Procedeu-se à recolha de um conjunto de indicadores e gráficos que seria relevante obter após a implementação do armazém de dados.

Foi apresentado o trabalho levado a cabo para determinar a importância e prioridade de implementação dos *data marts* a desenvolver.

Foram apresentadas as metodologias usadas na abordagem à fase de desenho do armazém de dados, nomeadamente, a construção de uma matriz de barramento dimensional como processo sistemático de determinação dos *data marts*, factos e dimensões que farão parte do futuro armazém de dados.

Foi ainda apresentada a metodologia dos quatro passos de Kimball (1998; 2002), usada para a especificação das tabelas de factos a implementar.

Os resultados desta fase foram finalmente expostos, tendo sido feita referência à matriz de barramento dimensional produzida e disponível no Anexo E, às tabelas de factos especificadas, exemplificadas e disponíveis no Anexo F e, por fim, foi feita referência às dimensões especificadas, exemplificadas e disponíveis para consulta no Anexo G.

6 Implementação

Neste capítulo serão apresentados os detalhes de implementação do armazém de dados. Inicia-se com a apresentação da tecnologia adotada, justificando as razões da sua escolha. Posteriormente, dependente da tecnologia usada, será apresentado o modelo físico. Por fim dar-se-á enfoque a questões particulares do desenvolvimento, relacionadas com a conceção do armazém e a extração, transformação e carregamento de dados.

6.1 Tecnologia

Para o desenvolvimento do armazém foram adotadas as seguintes ferramentas baseadas em tecnologia *Oracle*®: o sistema de gestão de base de dados *Oracle* para alojar o armazém de dados; a ferramenta *Oracle Warehouse Builder* como ferramenta de suporte à modelação, criação da base de dados e implementação das tarefas associadas ao seu carregamento; o *interface* de desenvolvimento *SQL Developer* como ferramenta de suporte à fase de extração e transformação dos dados; por último, para a apresentação de resultados foi utilizado o *Oracle Business Intelligence Discoverer*, que constitui um conjunto de ferramentas de interrogação, geração de relatórios e análise de dados, que vem integrado no pacote de ferramentas *Oracle*.

SISTEMA DE GESTÃO DE BASE DE DADOS *ORACLE*

O Sistema de Gestão de bases de Dados *Oracle* é a aplicação encarregue do armazenamento e gestão de dados, possuindo características que permitem gerir o seu acesso, integridade e segurança. Uma lista alargada das características do SGBD *Oracle* poderá ser consultada no capítulo dedicado à revisão bibliográfica, mais concretamente na secção dedicada às Infraestruturas de suporte à implementação de armazéns de dados, na página 22.

ORACLE WAREHOUSE BUILDER

O *Oracle Warehouse Builder* (OWB) é uma ferramenta que integra com o SGBD *Oracle*, suportando o modelo dimensional para a construção de armazéns de dados. Permite gerir e manter um meta-modelo do armazém de dados e ainda a criação física do referido armazém. Oferece ainda facilidades no que toca às operações de transformação e carregamento de dados, nomeadamente a geração automática de código que dá origem aos procedimentos de transformação e carregamento. Mais informações acerca do OWB poderão ser encontradas no

capítulo dedicado à revisão bibliográfica, mais concretamente na secção dedicada às Ferramentas de suporte à modelação de armazéns de dados, na página 27.

SQL DEVELOPER

O *SQL Developer* é uma *interface* de desenvolvimento que permite a ligação a bases de dados *Oracle* e o desenvolvimento de código PL/SQL. É especialmente útil como ferramenta de importação e visualização de dados, operações importantes na fase de extração de dados. Foi ainda utilizado para a criação de procedimentos PL/SQL que permitiram a transformação de dados, para que estes se apresentassem na forma mais adequada ao processo de carregamento.

ORACLE BUSINESS INTELLIGENCE DISCOVERER

O *Oracle Business Intelligence Discoverer* é a camada aplicacional encarregada de mediar a interação entre a base de dados e o utilizador final, possuindo características OLAP que permitem a exploração e análise de dados, sob a forma de tabelas e a apresentação de gráficos. Permite ainda a exportação de relatórios. Uma abordagem mais pormenorizada será apresentada no próximo capítulo, dedicado à apresentação de resultados.

6.1.1 Justificação da adoção das tecnologias usadas

Razões de natureza tecnológica e de natureza prática estão na origem das opções tomadas em favor da adoção de tecnologias *Oracle* para a implementação do armazém de dados.

As razões de ordem tecnológica são baseadas no trabalho de análise desenvolvido na fase de estudo do estado da arte e resultam da ponderação das características apresentadas pelos sistemas disponíveis. Como concluído na secção 2.5, verificou-se que qualquer uma das infraestruturas de suporte estudadas é tecnologicamente credível para a implementação de armazéns de dados, tendo sido especialmente valorizado o nível de integração das ferramentas *Oracle* em detrimento do aspeto negativo desta ser uma solução comercial, com custos associados. Facto que aliás é compensado pelas razões de ordem prática que serão evocadas a seguir.

Existem duas razões de ordem prática para a adoção destas tecnologias e que estão essencialmente associadas ao tempo disponível na implementação de um projeto desta natureza.

A primeira razão prende-se com o facto das soluções oferecidas pela *Oracle* existirem já na Faculdade, sendo a sua administração suportada pelo Centro de Informática. Desta forma houve uma poupança de tempo efetiva, eliminando à partida o tempo que eventualmente seria necessário para a instalação e administração de hipotéticas soluções que viessem a ser adotadas. Esta razão compensa também o aspeto negativo relacionado com facto das soluções *Oracle* serem soluções comerciais. Existindo elas na Faculdade, não existem custos acrescidos.

A segunda razão está relacionada com a experiência profissional e académica adquirida (Costa e Graça 2012), na utilização das soluções *Oracle* propostas, reduzindo desta forma o tempo associado à curva de aprendizagem de novas tecnologias.

6.2 Modelo físico

A implementação física do modelo lógico projetado está intimamente relacionada com a tecnologia adotada. Para soluções tecnológicas diferentes existem métodos de implementação diferentes, para os diversos conceitos do modelo dimensional.

No caso do *Oracle Warehouse Builder* esta questão é evidente na forma como é implementado o conceito de hierarquias e respetivos níveis. Da mesma forma, a inexistência de suporte a alguns conceitos do modelo dimensional, como por exemplo, suporte a dimensões degeneradas ou a tabelas de ponte, obriga a estratégias de implementação alternativas, à margem do modelo dimensional mantido no OWB.

O exemplo da criação de hierarquias é paradigmático da dicotomia existente entre modelo lógico e modelo físico.

A hierarquia de uma dimensão é uma estrutura lógica que usa níveis ordenados como meio de organização de dados. Descreve uma relação pai-filho entre um conjunto de níveis, com cardinalidade 1:n Oracle (2010, 4-6).

O OWB implementa hierarquias de dimensão recorrendo à criação automática de atributos na tabela, destinadas a registar chaves adicionais para cada nível da hierarquia. Paralelamente, são criados tuplos de controlo com chave negativa que possibilitam a ligação direta de factos a qualquer nível da hierarquia. Para cada nível da hierarquia de uma dimensão, uma linha de controlo será gerada, atuando assim como linha de ligação única para a tabela de factos Oracle (2010, 4-7).

Um exemplo da implementação de uma dimensão com hierarquia pode ser observado na Figura 6.2.1, onde são visíveis as chaves de dimensão (DIMENSION_KEY) e nível (CATEGORIA_ID e SUBCATEGORIA_ID), bem como os tuplos de controlo com chave negativa (os primeiros 4 tuplos), tudo isto implementado fisicamente, de modo a satisfazer as necessidades do modelo lógico.

DIMENSION_KEY	CATEGORIA_ID	CATEGORIA	DESCRICAO_CATEGORIA	SUBCATEGORIA_ID	SUBCATEGORIA	DESCRICAO_SUBCATEGORIA
-4	-4.4	ARQUIVO		(null) (null)		(null)
-3	-3.2	RECURSOS ELECTRÓNICOS		(null) (null)		(null)
-2	-2.3	SERVIÇOS		(null) (null)		(null)
-1	-1.1	COLEÇÕES DOS PISOS		(null) (null)		(null)
5	-1.1	COLEÇÕES DOS PISOS		5.1.1	Livros	
6	-1.1	COLEÇÕES DOS PISOS		6.1.2	Revistas	
7	-3.2	RECURSOS ELECTRÓNICOS		7.2.10	Patentes	
8	-3.2	RECURSOS ELECTRÓNICOS		8.2.11	Normas	
9	-3.2	RECURSOS ELECTRÓNICOS		9.2.12	Legislação	
10	-3.2	RECURSOS ELECTRÓNICOS		10.2.13	Obras de referência	
11	-3.2	RECURSOS ELECTRÓNICOS		11.2.14	Endnote	
12	-3.2	RECURSOS ELECTRÓNICOS		12.2.15	Outros recursos	
13	-3.2	RECURSOS ELECTRÓNICOS		13.2.3	Catálogo	
14	-3.2	RECURSOS ELECTRÓNICOS		14.2.4	Metalib	
15	-3.2	RECURSOS ELECTRÓNICOS		15.2.5	Portal	
16	-3.2	RECURSOS ELECTRÓNICOS		16.2.6	Revistas/artigos	
17	-3.2	RECURSOS ELECTRÓNICOS		17.2.7	Bases bibliográficas	
18	-3.2	RECURSOS ELECTRÓNICOS		18.2.8	eBooks	
19	-3.2	RECURSOS ELECTRÓNICOS		19.2.9	Teses e Dissertações	
20	-2.3	SERVIÇOS		20.3.1	Serviços de atendimento (Inscrição, Empr...	
21	-2.3	SERVIÇOS		21.3.2	Obtenção de Documentos (EIB, Aquisição d...	
22	-2.3	SERVIÇOS		22.3.3	Reserva de espaços	
23	-2.3	SERVIÇOS		23.3.4	Formação	
24	-4.4	ARQUIVO		24.4.1	Apoio presencial	
25	-4.4	ARQUIVO		25.4.2	Apoio remoto	

Figura 6.2.1: Exemplo de implementação de uma dimensão com hierarquias

Ainda relacionado com a temática do modelo físico, importa referir que o OWB, para além de proceder à manutenção do meta-modelo, gera e cria fisicamente todos os objetos de suporte ao armazém de dados, recorrendo para isso a diversas convenções internas pré-determinadas quanto à sintaxe a adotar relacionadas com os nomes a atribuir aos objetos. Apesar destes automatismos, o modelador tem sempre poder de intervenção. De uma maneira geral, na implementação deste armazém de dados, foram sempre respeitadas as regras pré-determinadas pelo OWB para a nomeação de objetos, onde se incluem, entre outras, as seguintes: tabelas físicas adotam os mesmos nomes dos objetos que as originam (dimensão ou cubo), seguidos do sufixo “_TAB”; tabelas físicas criadas para o registo de erros adotam os mesmos nomes dos objetos que as originam (dimensão ou cubo), seguido do sufixo “_ERR”; sequências adotam os mesmos nomes dos objetos que as originam, seguido do sufixo “_SEQ”; o nome de uma chave primária é formado concatenando o nome da tabela e do atributo associado, seguido do sufixo “_PK”.

Conclui-se esta secção dedicada ao modelo físico apresentando na Figura 6.2.2 um diagrama entidade-relação referente à implementação de uma estrela do armazém de dados, com o objetivo de tornar perceptível a implementação física e ao mesmo tempo tornar patente a maioria dos conceitos aqui apresentados, nomeadamente, as políticas de atribuição de nomes a objetos físicos, a implementação de hierarquias, níveis e seus atributos específicos, a utilização de dimensões degeneradas, de tabelas de ponte e até mesmo de implementação dimensional em floco-de-neve. Na figura a tabela de factos apresenta-se representada com fundo salmão, a tabela de ponte com fundo verde-claro e as dimensões com fundo branco.

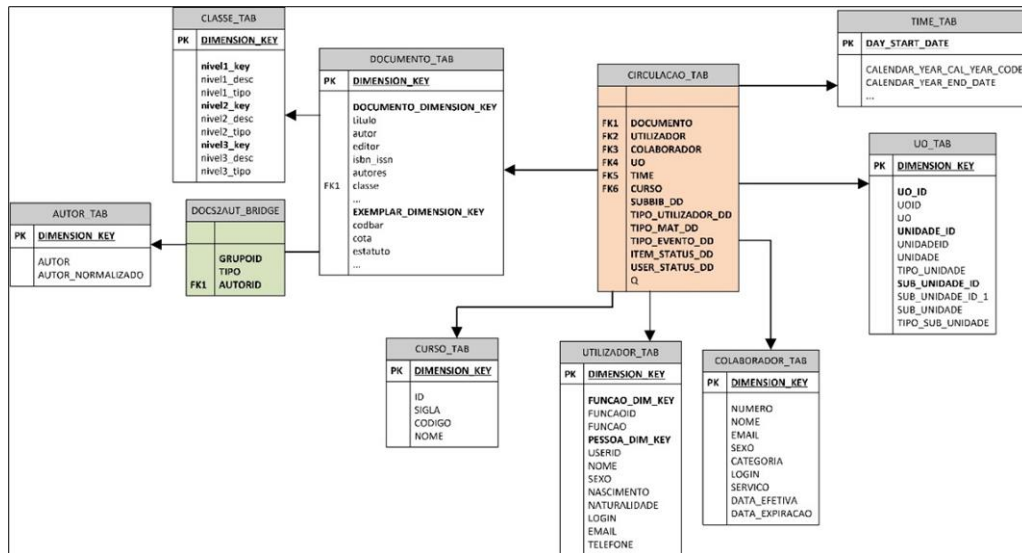


Figura 6.2.2: Implementação física da estrela “Circulação”

A dimensão “Documento” é um caso peculiar deste armazém de dados ao recorrer à implementação em floco de neve para representar a relação 1:n existente entre um documento e a sua classe. Por sua vez a dimensão “Classe” está hierarquicamente estruturada em três níveis, o que permite a navegação do nível mais específico até níveis mais gerais, possibilitando agrupar os documentos em grandes áreas. Por outro lado, a relação entre um documento e os seus autores é uma relação de m:n, implementada neste caso com recurso a uma tabela de ponte, ligando diretamente a dimensão “Documento” à dimensão “Autor”. A ligação de duas dimensões por via de uma tabela de ponte é uma opção pouco usual num modelo dimensional mas ainda assim válida e defendida por Kimball (2011). Outros exemplos de particularidades de implementação que influenciam o modelo físico poderão ser consultados com mais detalhe na secção 6.5.

A representação sob a forma de diagrama entidade-relação das restantes estrelas implementadas poderá ser consultada na íntegra no Anexo H.

6.3 Processo de implementação do armazém de dados

A implementação do armazém de dados é um processo iterativo e cíclico, tal como especificado no capítulo 3 dedicado à metodologia.

A Figura 6.3.1 abaixo apresentada sintetiza o método usado para a criação de estrelas durante o processo de implementação do armazém de dados.

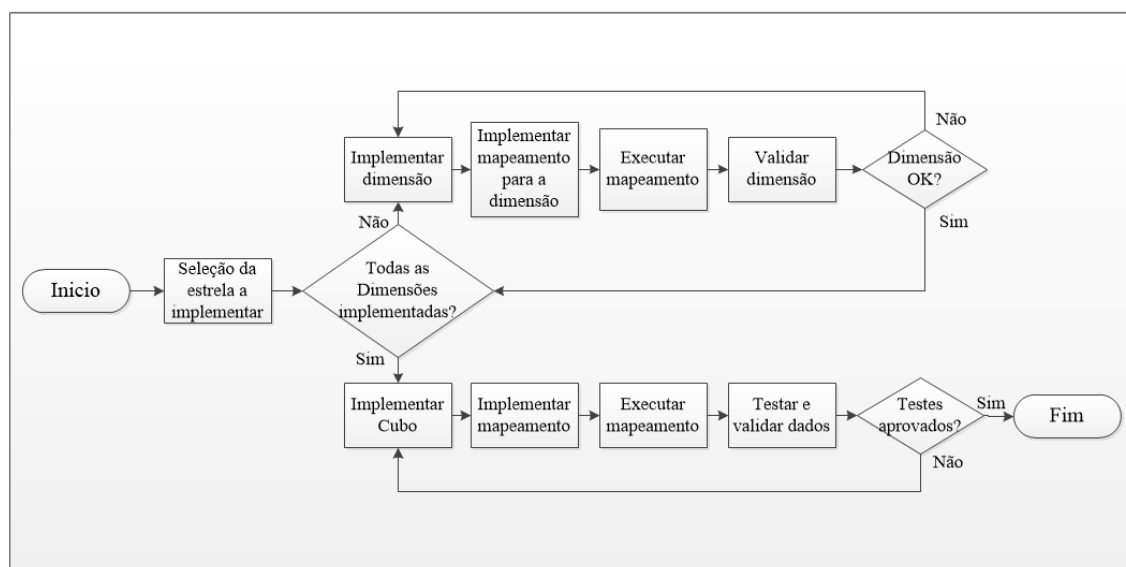


Figura 6.3.1: Diagrama de fluxo do processo de criação de estrelas

O processo inicia-se com a seleção da estrela a implementar. São validadas quais as dimensões projetadas para a estrela e quais se encontram já implementadas.

Caso alguma das dimensões associadas a essa estrela ainda não esteja implementada, procede-se à implementação dessa dimensão e à criação do mapeamento que a irá preencher. Executa-se o mapeamento para preencher a dimensão. Após validação positiva dos dados mapeados, verifica-se novamente a existência de mais dimensões por implementar. Caso existam, voltamos a repetir o processo de criação de dimensões até que não exista mais nenhuma por implementar, associada à estrela que está a ser desenvolvida.

Caso todas as dimensões associadas à estrela estejam já implementadas, procede-se então à criação do cubo da estrela, à criação do mapeamento que irá preencher a tabela de factos. Executa-se o mapeamento de dados que preenche o cubo, sendo posteriormente efetuados testes de validação dos dados carregados. Caso a avaliação seja positiva, é selecionada outra estrela para implementar, repetindo-se novamente todos os passos anteriores. Caso seja negativa, o processo de implementação da estrela repete-se, procedendo-se às medidas corretivas necessárias para a obtenção de uma avaliação positiva.

6.3.1 Criação de uma dimensão

A criação de objetos no OWB é concretizada através do recurso a assistentes, que ao longo de vários passos interrogam o utilizador sobre as configurações indispensáveis à sua criação.

O assistente de criação de dimensões interroga o utilizador sobre as seguintes configurações: nome da dimensão; tipo de armazenamento (ROLAP ou MOLAP); atributos da dimensão (incluindo nome, tipo, tamanho, precisão e tipo de identificador); níveis na hierarquia por omissão; quais os atributos associados a cada nível e tipo de registo de histórico (SCD I, SCD II ou SCD III).

Após a criação da dimensão, para além de ser possível ajustar as configurações referidas anteriormente, é ainda possível configurar aspetos relacionados com as hierarquias e com a gestão efetuada ao nível do carregamento de dados, nomeadamente, se se pretende a criação de uma tabela para o registo de erros de carregamento e ainda a forma como são tratados os dados cuja chave é inválida ou nula.

A Figura 6.3.2, abaixo apresentada, pretende ilustrar o assistente de configuração de uma dimensão. No exemplo apresentado, podemos verificar os seus separadores, dando ênfase ao separador onde é possível observar resumidamente a convivência entre o modelo lógico e o modelo físico implementado, bem como muitas das suas características, como sejam: as chaves da dimensão, os atributos, as hierarquias e respetivos níveis, a ligação à tabela física, o nome da tabela física, os nomes dos atributos na tabela física e as chaves criadas na tabela física.

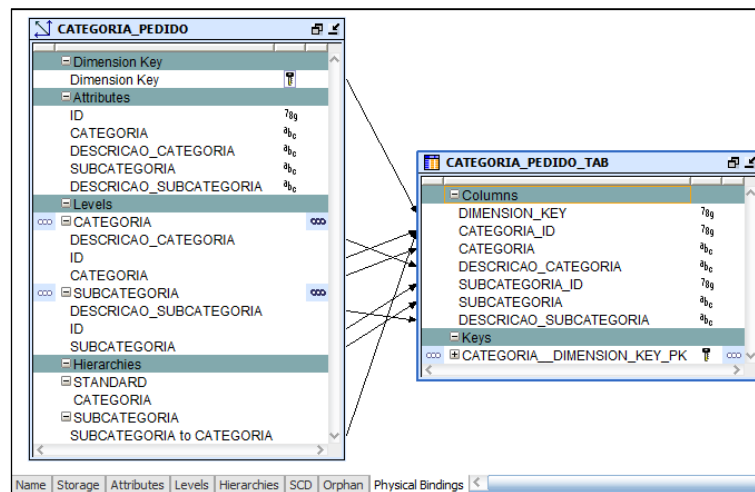


Figura 6.3.2: Assistente de configuração de dimensões. Ligação entre modelo lógico e modelo físico.

6.3.2 Criação de um cubo

A criação de um cubo, à semelhança da criação de uma dimensão, é concretizada recorrendo a um assistente de configuração. Existem quatro passos indispensáveis para recolher a informação necessária à sua implementação: especificação do nome do cubo; do tipo de armazenamento (ROLAP ou MOLAP); das dimensões que participam na estrela e as medidas que o cubo pretende capturar (incluindo nome, tipo, tamanho, precisão e escala).

O comportamento do cubo relativamente ao tipo de agregação calculada e ao tipo de gestão de dados com chaves nulas ou inválidas é ajustado posteriormente.

No momento do carregamento, para além dos dados carregados, é possível instruir o sistema para, paralelamente, calcular agregações relacionadas com esses dados. O emprego destes métodos visa a otimização da resposta do sistema no momento da recolha de resultados. Existem vários métodos segundo os quais os dados podem ser agregados e que dependem do tipo de dados em causa e do propósito para o qual esses dados são utilizados. Vários métodos de agregação estão ao dispor do utilizador no OWB, desde o método de soma de medidas, ao cálculo de médias, mínimos, máximos, passando pela determinação do primeiro ou do último, ou ainda pela contagem de factos, terminando com a opção de não existência de qualquer tipo de agregação. Na implementação deste armazém de dados foi sempre utilizado o somatório como método de agregação preferencial.

Relativamente ao comportamento do cubo face a dados com chaves inválidas ou nulas, à semelhança das dimensões, é possível modelar o seu comportamento de forma a usar um registo de dimensão por defeito, rejeitar o registo ou simplesmente não fazer qualquer espécie de manutenção. É também possível recorrer a uma tabela para registar os erros de carregamento.

A Figura 6.3.3, abaixo apresentada, pretende ilustrar o assistente de configuração de um cubo. No exemplo apresentado, podemos verificar os seus separadores, dando ênfase ao separador onde é possível observar resumidamente a convivência entre o modelo lógico e o modelo físico implementado, bem como muitas das suas características, como sejam: as dimensões participantes no cubo, as medidas do cubo, a ligação à tabela física, o nome desta tabela, os nomes dos seus atributos e a especificação das chaves.

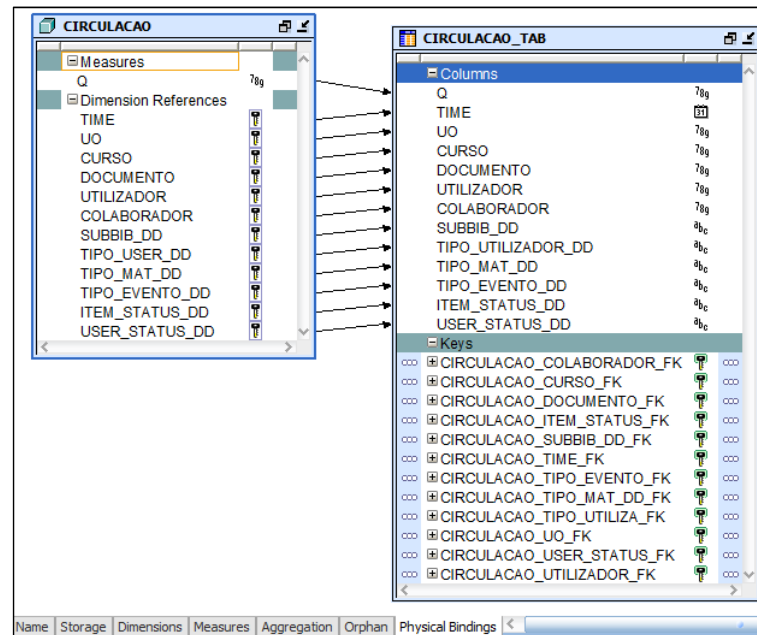


Figura 6.3.3: Assistente de configuração de um cubo. Ligação entre modelo lógico e modelo físico.

6.4 Extração, transformação e carregamento

A extração, transformação e carregamento são os processos através dos quais são obtidos e modelados os dados que alimentam o armazém. Constituem as tarefas mais complexas e críticas de todo o processo de construção de um armazém de dados. A potencial diversidade de origem de dados, quer ao nível dos sistemas operacionais envolvidos e respetivas tecnologias associadas, quer ao nível do formato dos próprios dados exige alguma multidisciplinaridade e por vezes a aquisição de novas competências.

Regra geral, neste caso concreto, tentou sempre obter-se os dados de forma a que estes pudessem ser importados para tabelas *Oracle* do servidor de alunos, com o auxílio do *SQL Developer*, para posteriormente se proceder às transformações que eventualmente fossem necessárias.

Embora a conceção do modelo dimensional para a implementação do armazém apresentasse como pré-requisito a existência de dados capazes de cobrir todas as expectativas, a verdade é que durante a fase de implementação depressa se concluiu que em alguns casos estes não correspondiam ao esperado, obrigando ao redesenho de estrelas como forma de colmatar as lacunas de dados existente e assim poder satisfazer, de alguma maneira, os indicadores previstos.

Foram identificadas duas situações nas quais os dados ideais e a sua realidade divergiam:

- No caso em que a estrela que responde a determinado indicador tem prevista determinada dimensão, mas nos dados existentes essa dimensão não existe;
- No caso onde existem divergências de granularidade entre o projetado e os dados disponíveis;

Um exemplo prático desta divergência pode ser encontrado na estrela que pretende medir as consultas ao arquivo, acompanhadas por colaboradores.

	A	B	C	D	E	F	G	H	I	J
1	ID	número	nome	código	sigla	nome curso	ano letivo	ano curricular	estado	Mobilidade
2	718124	200807129	Ana		API	O Processo de Inovação	2008		1 A Frequentar	Não
3	727288	199102377	Silvia		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
4	727278	199503392	Elisabe		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
5	727280	200103623	Antóni		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
6	727283	200200366	Diana		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
7	729030	200406340	Bruno		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
8	727285	200502097	Nuno		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
9	727281	200601445	Silvia		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
10	727282	201101845	Raquel		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
11	727284	201104301	Rui		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
12	727286	201105459	Paula		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
13	727279	201107954	Cidália		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
14	727287	201107955	Sandra		BRE	Beneficiação - Reabilitação de Estradas	2011		1 A Frequentar	N
15	830098	200102291	Diana		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
16	833836	200102849	Vitor		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
17	833817	200305451	João		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
18	830059	201208729	Huseyi		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
19	833814	201209208	Esra		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
20	833815	201209209	Hüseyi		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
21	833816	201209210	Hüseyi		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
22	833818	201209211	Kazim		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
23	833819	201209212	Muham		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
24	833820	201209213	Osman		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
25	833821	201209214	Özge		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
26	833834	201209215	Süleym		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
27	833835	201209216	Ugur		CERS	Comportamento de Estruturas em Regiões	2012		1 A Frequentar	N
28	707022	198400716	Sofia	3117	CPGCE	Especialização em Construção de Edifícios	2006		1 A Frequentar	N
29	707022	198400716	Sofia	3117	CPGCE	Especialização em Construção de Edifícios	2006		1 Concluído	N

Figura 6.4.2: Exemplo de origem de dados. Lista de estudantes e seu percurso académico.

6.4.2 Transformação

O objetivo desta fase é modelar os dados originais, colocando-os num formato compatível com o lugar a que se destinam no armazém de dados, seja ele uma dimensão ou uma tabela de factos de uma estrela.

No processo de criação do armazém de dados do SDI, a transformação de dados recorreu sobretudo a operações realizadas com o *SQL Developer*, através de *scripts PL/SQL*, ou em casos extremos, a operações de modelação de dados em ficheiros *Excel*.

Dois exemplos ilustrativos deste tipo de procedimento são abaixo apresentados.

No primeiro caso, a Figura 6.4.3, mostra um excerto de uma folha *Excel* com dados relacionados com o processo de criação da dimensão “Classe”, a classificação *ISAD* utilizada no arquivo, onde os dados obtidos a partir de um relatório em formato *Excel* possuía colunas com caracteres de controlo (CRLF), causando problemas de importação. No caso apresentado ilustra-se a utilização da função *CLEAN* do *Excel* para corrigir esse problema.

E18931					:	X	✓	fx	=CLEAN([@Classificação])
	id	Código	Entidade Produtora	Classificação	Column1				
18931	54128	SECPOS320	Secção de Pós-Graduação (2004 ????)	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/				
18932	54194	SECPOS321	Secção de Pós-Graduação (2004 ????)	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/				
18933	54466	SECPOS322	Secção de Pós-Graduação (2004 ????)	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/				
18934	54772	SECPOS323	Secção de Pós-Graduação (2004 ????)	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/	PT-UP- FEUP/FEUP/CNSDIR/SRVAC/D/IVPOSEC/SECPOS2/17500/				

Figura 6.4.3: Exemplo de operação de transformação realizada numa folha *Excel*.

No segundo caso, a Figura 6.4.4 mostra um excerto de um *script PL/SQL*. Este *script* foi utilizado para corrigir os números de estudante, substituindo-os pelos números novos que começaram a ser utilizados na Universidade do Porto no presente ano letivo. Os dados referentes aos factos que registam as transações relacionadas com a circulação na Biblioteca (empréstimos, devoluções, renovações, etc...) foram recuperados de *backups* antigos, foram importados para uma tabela *Oracle*, sofreram diversas operações de ajuste (*trim*, *pad*, concatenação, etc...), foram enriquecidos com referências chave do curso ou do departamento do utilizador e por fim o procedimento *PL/SQL* atuou no campo de identificação do utilizador para que o número antigo fosse substituído pelo novo número (a chave de utilizador na dimensão respetiva).

```

PROCEDURE CORRIGE_Z35 AS
CURSOR OLDSTUDENTNUMBERSCURSOR IS
SELECT * FROM NEWNALUNOS WHERE VELHOLD IN (SELECT DISTINCT TRIM(USERID) FROM EXT_EUP50Z35_2011) ;
CURRENTOLDID NEWNALUNOS%ROWTYPE;
BEGIN
DBMS_OUTPUT.ENABLE(NULL);
OPEN OLDSTUDENTNUMBERSCURSOR;
LOOP
FETCH OLDSTUDENTNUMBERSCURSOR INTO CURRENTOLDID;
EXIT WHEN OLDSTUDENTNUMBERSCURSOR%NOTFOUND;
UPDATE EXT_EUP50Z35_2011 SET USERID = CURRENTOLDID.NOVOID WHERE USERID LIKE CURRENTOLDID.VELHOLD||'&';
DBMS_OUTPUT.PUT_LINE('DEBUG]A ATUALIZAR ID ANTIGO: '||CURRENTOLDID.VELHOLD);
COMMIT;
END LOOP;
CLOSE OLDSTUDENTNUMBERSCURSOR;
EXCEPTION
WHEN OTHERS THEN
DBMS_OUTPUT.PUT_LINE('Operation failed '||'SQLCODE: '||SQLCODE);
DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
END CORRIGE_Z35;

```

Figura 6.4.4: Exemplo de *script* utilizado nas operações de transformação.

Uma nota final apenas para mencionar que o processo de transformação de dados poderia também ser realizado recorrendo à técnica de mapeamento de dados, coadjuvado pela vasta lista de operadores disponíveis no OWB. Tal não sucedeu, privilegiando-se o emprego de mapeamentos unicamente na fase de carregamento de dados.

Para o OWB, um mapeamento descreve uma série de operações de extração, transformação e carregamento de dados, oferecendo uma representação visual do fluxo e das operações sobre os dados.

Os operadores são o elemento gráfico básico de um mapeamento. Através dos operadores são representadas a origem e destino no fluxo de dados, bem como as operações de transformação realizadas pelo caminho Oracle (2010, 5-1) .

6.4.3 Carregamento

Após a conclusão da fase de transformação, o carregamento de dados foi essencialmente realizado recorrendo a mapeamentos criados no OWB. Para que tal fosse possível, o meta-modelo das tabelas com as origens de dados, já transformadas, foram importadas para o OWB, de forma a poderem participar no mapeamento de dados criados.

A Figura 6.4.5, abaixo apresentada mostra o exemplo do mapeamento criado para carregar dados na tabela de factos da estrela “Documentos novos”, que pretende medir o número de documentos novos registados nos sistemas do serviço.

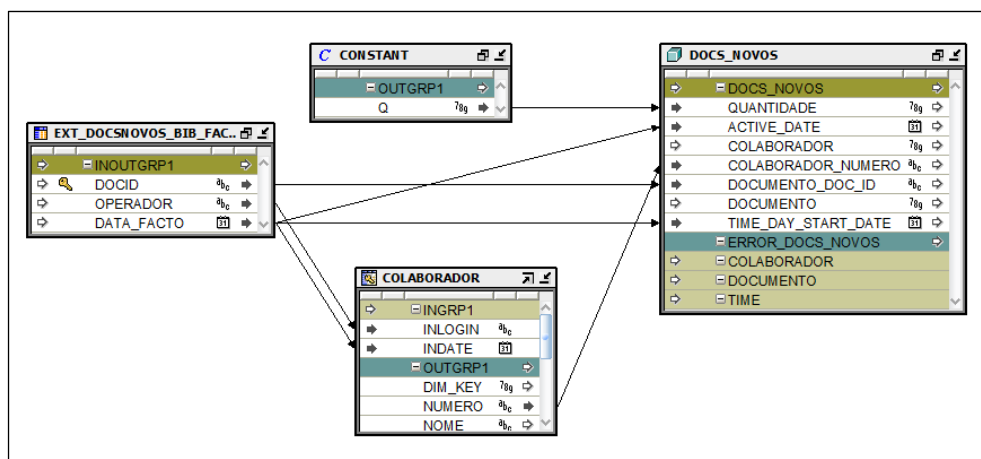


Figura 6.4.5: Exemplo de um mapeamento para carregamento de dados

No exemplo apresentado pode verificar-se, à esquerda, a tabela externa que representa a origem de dados, à direita, a representação do cubo a carregar com dados e ao centro o emprego de operadores para auxiliar o preenchimento de dados. No caso específico, o emprego dos operadores “Constant” e “Lookup table”. O primeiro operador serve para preencher a medida do cubo com um valor constante (1). O segundo operador foi necessário para descobrir na dimensão “Colaborador” qual a chave do operador envolvido, dado a origem de dados apenas apresentar o login. Desta forma, o operador “Lookup table” consulta uma tabela de acordo com os dados de entrada, devolvendo o registo de colaborador apropriado.

A Figura 6.4.6 representa um excerto do código gerado automaticamente pelo OWB, fruto da compilação do mapeamento apresentado na Figura 6.4.5. São visíveis ainda à esquerda os diversos procedimentos e funções que compõem o pacote compilado.

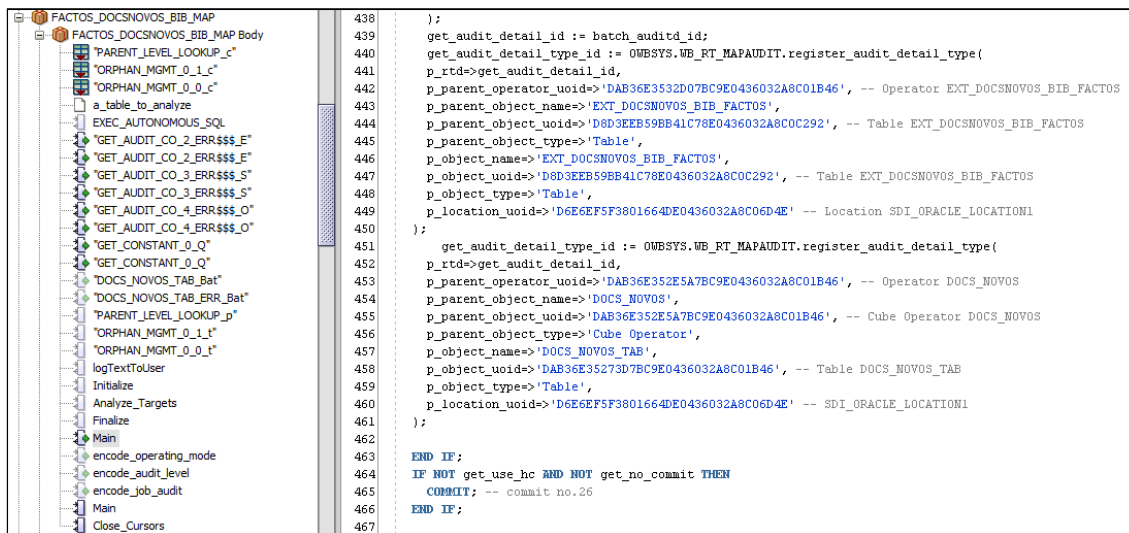


Figura 6.4.6: Vista sobre pacote PL/SQL gerado automaticamente para mapeamento

6.5 Particularidades de implementação

Pretende-se nesta secção deixar registados alguns aspetos da implementação que se destacam no desenvolvimento global do armazém.

Essas particularidades resultam em parte da falta de cobertura do OWB a alguns conceitos do modelo dimensional, como por exemplo, suporte a dimensões degeneradas ou a tabelas de ponte. Outras particularidades resultam de opções específicas de implementação.

6.5.1 Implementação de tabelas de factos com dimensões degeneradas

Uma dimensão degenerada é uma dimensão que não tem a sua própria tabela de dimensão, é representada por um valor único, sem atributos adicionais e que por este motivo fica registada na própria tabela de factos (Kimball e Ross 2002).

A implementação de dimensões degeneradas é um dos casos de conceitos do modelo dimensional que não é coberto no OWB. Devido a esta circunstância, a sua implementação deverá ser realizada recorrendo a técnicas alternativas.

Para criar dimensões degeneradas para o armazém de dados, foi seguida a metodologia sugerida por Allan (2010) e que se resume sinteticamente a criar a dimensão degenerada, utilizando o habitual assistente de criação de dimensões do OWB, tendo o cuidado de apagar a

chave da dimensão e a hierarquia criada por omissão. A seguir, deve ser criada a tabela física que suporta essa dimensão, embora não necessite de receber dados. Finalmente o cubo onde a referida dimensão degenerada participa deve ser criada, tendo o cuidado de configurar a sua tabela de factos para que na altura da sua criação não implemente nenhuma chave estrangeira para a dimensão degenerada. A Figura 6.5.1 ilustra essa configuração para a criação da tabela de factos do cubo “Multas”.

Property	DEFAULT_CONFIGURATION
<ul style="list-style-type: none"> MULTAS_TAB <ul style="list-style-type: none"> Generation Comments Change Data Capture Columns Error Table Foreign Keys <ul style="list-style-type: none"> MULTAS_COLABORADOR_FK MULTAS_DATA_I_FK MULTAS_DATA_P_FK MULTAS_DOCUMENTO_FK MULTAS_ESTADO_TRANSA_FK MULTAS_NUMERO_RECIBO_FK Creation Method Identification 	
Deployable	false
<ul style="list-style-type: none"> <ul style="list-style-type: none"> MULTAS_TIPO_TRANSACA_FK MULTAS_UTILIZADOR_FK Identification Indexes Parallel Partition Parameters Performance Parameters Physical Attributes Storage Space 	

Figura 6.5.1: Configuração da tabela de um cubo para suportar dimensões degeneradas

Desta forma é possível usar o mapeamento de dados para preencher a tabela de factos através do operador de cubo, sem correr o risco da ocorrência de erros devido ao facto de existir uma dimensão degenerada.

6.5.2 Implementação de tabelas de ponte

Outro exemplo de conceitos do modelo dimensional não coberto pelo OWB é a implementação de tabelas de ponte.

Uma tabela de ponte é uma tabela com uma chave múltipla que é usada para capturar uma relação de muitos-para-muitos que não pode ser acomodada pela granularidade natural de uma tabela de factos ou de uma dimensão. É também conhecida por tabela associativa, servindo normalmente para ligar factos a vários valores de uma dimensão (Kimball e Ross 2002).

A Figura 6.5.2, em baixo, apresenta o diagrama entidade-relação da estrela “Indexação” e pretende ilustrar a implementação de duas tabelas de ponte. A primeira (*INDEX_BRIDGE*) segue uma abordagem mais tradicional ligando a tabela de factos à dimensão “Assunto”. A segunda (*DOCS2AUT_BRIDGE*) segue uma implementação não tão usual mas ainda assim válida, suportada por Kimball (2011), na qual duas dimensões são ligadas recorrendo a uma tabela de ponte.

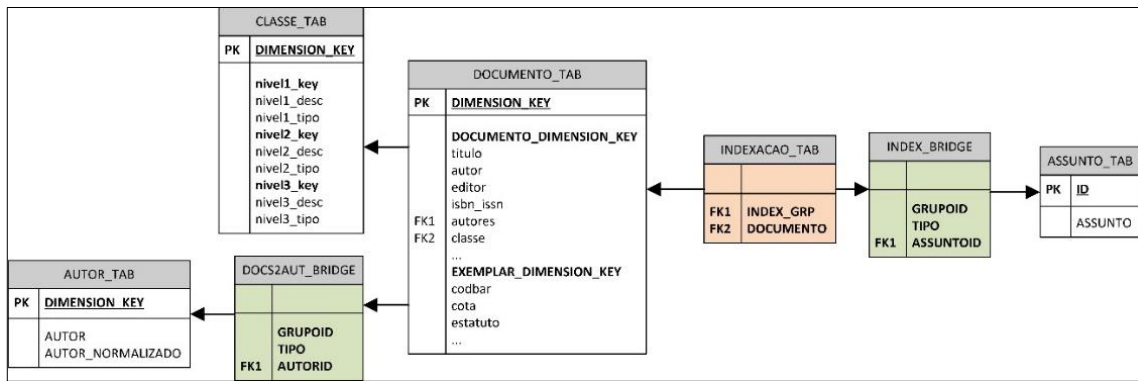


Figura 6.5.2: Exemplo de implementação de tabelas-de-ponte

Para determinar os grupos distintos de autores e assuntos dos documentos, de forma a poder estabelecer identificadores chave para esses grupos e preencher as tabelas de ponte, procedeu-se à criação de uma função para a agregação de dados, documentada por Oracle (2001), que permitiu extrair todos os grupos distintos, marcá-los com um identificador de grupo e por fim proceder à criação da tabela de ponte, estabelecendo a ligação entre as dimensões e a tabela de factos.

O código que serviu de base para a definição da função de agregação é apresentado em baixo. Primeiro, na Figura 6.5.3, a especificação de um tipo de dados definido pelo utilizador.

```
create or replace type string_agg_type as object
(
    total varchar2(4000),

    static function
        ODCIAggregateInitialize(sctx IN OUT string_agg_type)
        return number,

    member function
        ODCIAggregateIterate(self IN OUT string_agg_type,
                             value IN varchar2)
        return number,

    member function
        ODCIAggregateTerminate(self IN string_agg_type,
                               returnValue OUT varchar2,
                               flags IN number)
        return number,

    member function
        ODCIAggregateMerge(self IN OUT string_agg_type,
                            ctx2 IN string_agg_type)
        RETURN NUMBER
);

create or replace type body string_agg_type
is
    static function ODCIAggregateInitialize(sctx IN OUT string_agg_type)
    return number
    is
    begin
        sctx := string_agg_type( null );
        return ODCIConst.Success;
    end;

    member function ODCIAggregateIterate(self IN OUT string_agg_type,
                                         value IN varchar2)
    return number
    is
    begin
        self.total := self.total || ',' || value;
        return ODCIConst.Success;
    end;

    member function ODCIAggregateTerminate(self IN string_agg_type,
                                           returnValue OUT varchar2,
                                           flags IN number)
    return number
    is
    begin
        returnValue := ltrim(self.total, ',');
        return ODCIConst.Success;
    end;

    member function ODCIAggregateMerge(self IN OUT string_agg_type,
                                       ctx2 IN string_agg_type)
    return number
    is
    begin
        self.total := self.total || ctx2.total;
        return ODCIConst.Success;
    end;
end;
```

Figura 6.5.3: Tipo de dados definido pelo utilizador, usado na função de agregação

Na Figura 6.5.4 o corpo da função que usa o tipo de dados definido anteriormente.

```
create or replace
FUNCTION stragg(input varchar2)
RETURN VARCHAR2
PARALLEL_ENABLE AGGREGATE USING string_agg_type;
```

Figura 6.5.4: Função STRAGG

Por fim, a Figura 6.5.5 mostra um exemplo de uma interrogação SQL efetuada ao sistema operacional para a extração dos grupos de autores, que foram importados posteriormente para uma tabela auxiliar.

1	SELECT DOCID, STRAGG(AUTOR) FROM AUTORES GROUP BY DOCID;	
Query Result x		
Fetched 150 rows in 1,336 seconds		
	DOCID	STRAGG(AUTOR)
106	000000121	Shurcliff, William A.
107	000000122	Franklin Research Center
108	000000123	Wright, David
109	000000135	Kinloch, A. J.
110	000000136	Sadek, Mohammed Mohamed
111	000000137	Nichols, R. W.
112	000000138	Solomon, H. D.,ASTM.,Symposium on Low Cicle Fatigue,
113	000000139	Atkinson, P.
114	000000140	Anthony, Derek M.
115	000000141	Harris, John Noel
116	000000142	Sully, F. K.
117	000000143	Viswanathan, Ramaswamy
118	000000144	Wulpi, Donald J.
119	000000145	Andrejko, Dennis A.,National Passive Solar Conference,,Hayes, John
120	000000146	Andrejko, Dennis A.,Hayes, John,Passive Solar Conference,
121	000000147	Portugal.,Seminário
122	000000148	Colóquio sobre Realidades e Perspectivas da Investigação Científica no Ensino Superior,,Portugal.

Figura 6.5.5: Extração de grupos de autores

A Figura 6.5.6 exemplifica a técnica usada para o preenchimento de um tuplo de uma tabela-de-ponte e a ligação desta às tabelas de dimensão e/ou factos.

Na imagem é visível a presença de três tabelas auxiliares, representadas com a cor verde, necessárias ao preenchimento da tabela-de-ponte. No passo 1, um nome de um autor é retirado da tabela auxiliar (*EXT_AUTORES_DOCS*) e é pesquisado na tabela da dimensão de autores.

Uma vez encontrado esse autor, nos passos 2 e 3, a respetiva chave de dimensão é colocada num tuplo da tabela-de-ponte, acompanhada do tipo de autor.

A determinação do identificador do grupo de autores que irá completar o tuplo inicia-se com o passo 4, onde a chave de negócio do documento é verificado contra uma segunda tabela auxiliar (*EXT_DOCS_STRAGG*), constituída por uma linha para cada documento, onde se pode encontrar o conjunto de autores agregados com a função “*STRAGG*”, explicada anteriormente.

O passo 5 procura o valor com o conjunto de autores agregados na terceira e última tabela auxiliar (*EXT_AUTORES_GRUPOS*), onde se estabelece a relação entre esse conjunto e a chave de identificação do grupo de autores.

O preenchimento do tuplo na tabela-de-ponte é concluído no passo 6 com a inserção do valor da chave de grupo, juntamente com o valor do tipo de autor e a chave de dimensão do autor, determinados nos passos anteriores.

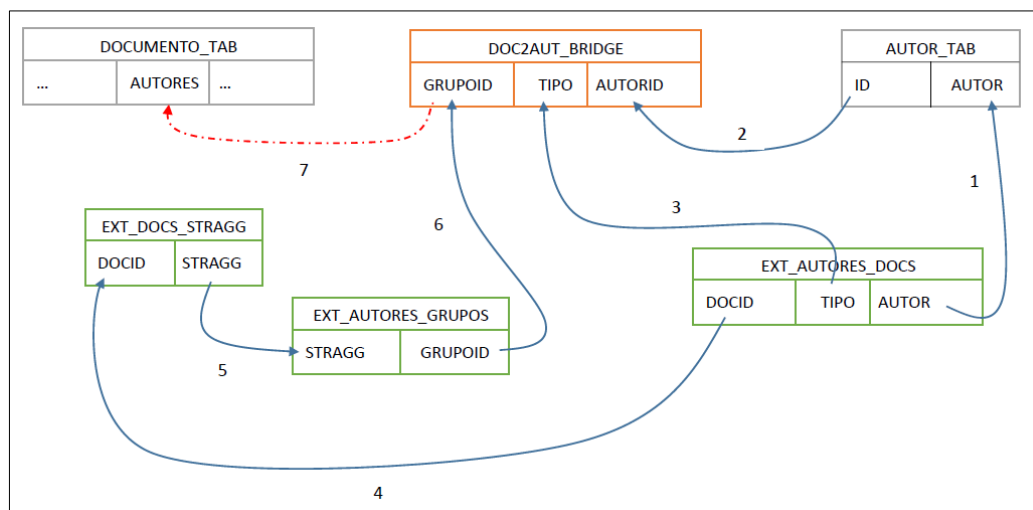


Figura 6.5.6: Processo de preenchimento de uma tabela de ponte

O passo 7, representado pela linha tracejada vermelha, é processado posteriormente e tem como objetivo o preenchimento do atributo destinado a armazenar a chave do grupo de autores do documento.

O código responsável pelo processamento descrito anteriormente pode ser encontrado no anexo I.1.

6.5.3 Tabelas de factos que usam *default dimension keys*

Uma funcionalidade particular permitida pelo OWB é a possibilidade das tabelas de factos poderem receber como valor de chave de dimensão uma chave por omissão, quando no momento do carregamento de dados, para uma dada dimensão, são detetadas chaves inválidas ou nulas.

Este mecanismo permite implementar tabelas de factos onde existem duas dimensões mutuamente exclusivas, isto é, a presença de uma impossibilita a presença da outra e por isso uma delas apresenta sempre uma chave nula.

A Figura 6.5.7, abaixo apresentada, representa a implementação física da estrela “Circulação” onde esta solução foi implementada.

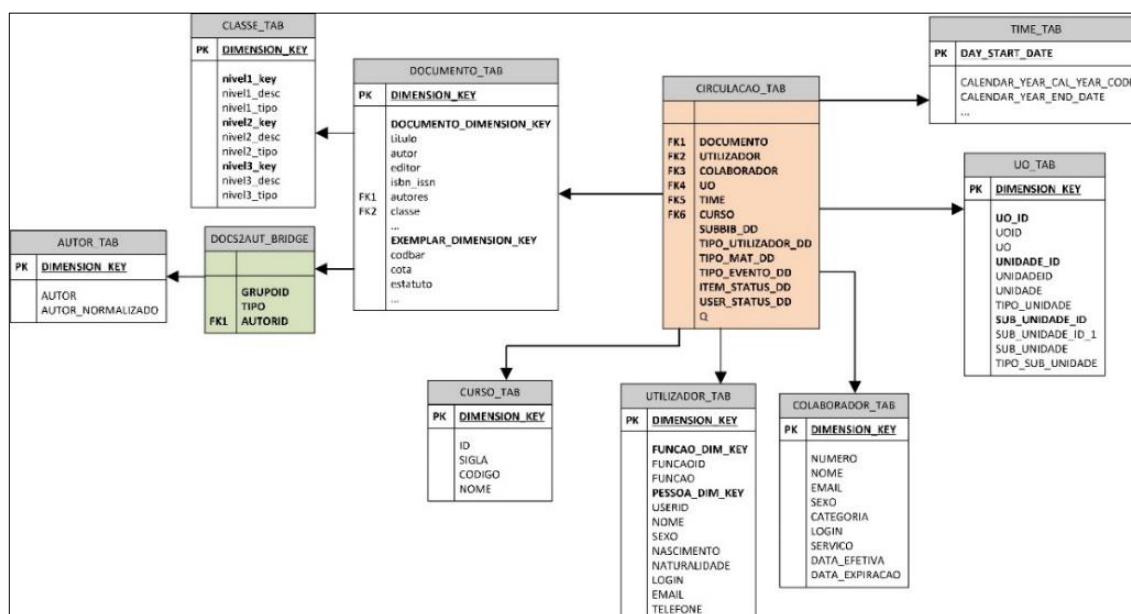


Figura 6.5.7: Diagrama entidade-relação com a implementação da estrela “Circulação”

A estrela “Circulação” está encarregada da captura de eventos de empréstimo, devolução, renovação, reserva e consulta de documentos. A presença da dimensão Curso e da dimensão da Unidade Orgânica (UO) possibilita determinar a afiliação do utilizador. Porém, um utilizador pertence a um Curso, caso seja do tipo “estudante” ou pertence a um Departamento ou Serviço, caso seja “funcionário” ou “docente”. Nunca pertence aos dois ao mesmo tempo. Com este mecanismo a coabitação das duas dimensões foi possível na mesma estrela.

6.5.4 *Slowly Changing Dimensions*

Slowly Changing Dimension (SCD) é o comportamento das linhas de uma dimensão quando estas mudam gradual ou ocasionalmente com o decorrer do tempo. Existem três tipos de SCD. A SCD do tipo 1 é a dimensão em que os valores são reescritos quando os atributos de uma dimensão são alterados. A SCD do tipo 2 é a dimensão onde uma nova linha é criada quando um valor de um atributo muda. A SCD do tipo 3 é a dimensão onde uma coluna alternativa para o valor antigo é criada, sempre que determinado atributo muda (Kimball e Ross 2002).

Na implementação do armazém de dados para o SDI foi utilizada a implementação de SCD do tipo 2 para a dimensão “Colaborador”. Para alguns dos colaboradores aí registados foi detetada a existência de alterações ao longo do tempo no atributo “login”.

A criação da dimensão com SCD tipo 2 obrigou à definição de dois novos atributos que servem para determinar a validade do registo e determinam qual o registo válido para as operações de carregamento de dados.

A Figura 6.5.8 apresenta a configuração do mecanismo de SCD do tipo 2 para a dimensão “Colaborador”, onde é patente a declaração dos atributos da dimensão que despoletam a criação de um novo registo sempre que se verifica uma alteração e dos atributos onde as datas de validade são registadas. As chaves de utilizador não podem fazer parte deste mecanismo.

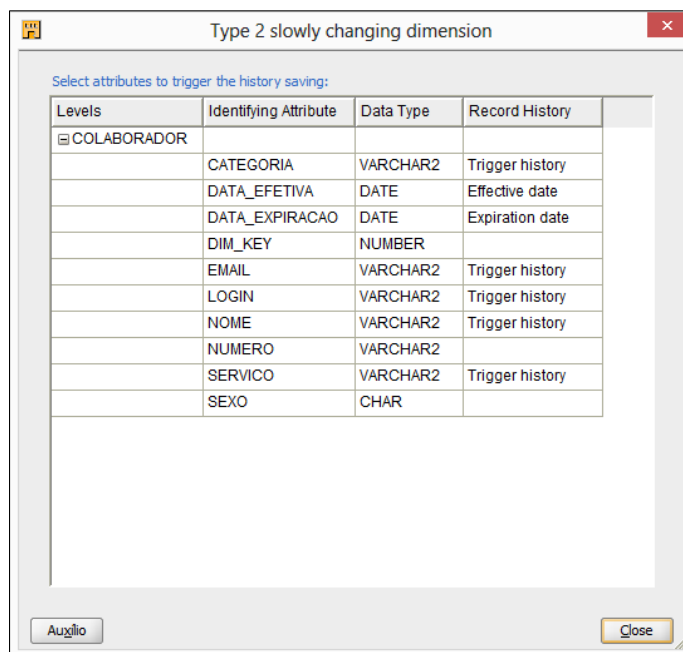


Figura 6.5.8: Configurações SCD 2 para a dimensão colaborador

6.6 Conclusões

Neste capítulo foi abordada a temática da implementação, apresentando-se inicialmente a tecnologia adotada para o desenvolvimento do armazém de dados, tendo o cuidado de justificar as opções tomadas neste domínio.

O protótipo efetivamente implementado corresponde a três data marts, num total de cerca de duas dezenas de estrelas, com dados entre 2007 e 2012.

Posteriormente, foram apresentadas as particularidades do modelo físico implementado, procurando evidenciar, através de exemplos, os casos onde o modelo físico se distancia do modelo lógico projetado.

Para culminar a secção dedicada ao modelo físico, foi apresentado o exemplo de um diagrama entidade-relação referente à implementação de uma estrela do armazém de dados e remetendo para o Anexo H a consulta dos restantes diagramas.

A descrição do processo de implementação do armazém de dados foi abordada a seguir, iniciando-se com a apresentação de um diagrama de fluxo que sintetiza o processo de criação e finalizando com a descrição do processo de criação das unidades básicas de um armazém de dados: a dimensão e o cubo.

Foi descrito o processo de extração, transformação e carregamento de dados adotado na implementação, procurando sempre apresentar exemplos que evidenciem as opções tomadas neste contexto.

O capítulo finaliza com a apresentação de particularidades de implementação, que de algum modo se diferenciam no desenvolvimento global do armazém. No rol de particularidades apresentadas podem ser encontrados pormenores relacionados com a implementação: de dimensões degeneradas; de tabelas-de-ponte; de tabelas de factos que usam *default dimension keys* e da implementação de dimensões que usam *slowly changing dimension* (tipo 2).

No decurso do processo de implementação, verificou-se que as tecnologias de desenvolvimento escolhidas serviram o seu propósito de forma eficiente, não obstante existirem alguns obstáculos cuja superação foi conseguida recorrendo a técnicas alternativas, levando desta forma a uma readaptação da metodologia de desenvolvimento.

Da disparidade existente entre o modelo idealizado e a verdadeira natureza dos dados, foi necessário empreender ações tendentes a procurar oferecer soluções viáveis em detrimento das soluções inicialmente idealizadas. A diversidade de fontes de dados, de agentes responsáveis pela sua recolha, a diferença de granularidade dos dados esperada e a encontrada, determinou a frequente utilização destas soluções.

7 Apresentação de Resultados

No seguimento da metodologia adotada para o desenvolvimento desta dissertação, na qual se segue uma abordagem de estudo de caso e se privilegia um desenvolvimento iterativo, após a implementação de alguns *data marts*, é chegada a fase de extração de indicadores e apresentação de resultados à Direção de Serviços, num processo que se pretende que seja mais uma etapa no processo de aceitação do trabalho desenvolvido.

Deste modo, procedeu-se ao desenvolvimento de indicadores e gráficos, baseados nos dados já carregados no armazém de dados, sendo este trabalho e a subsequente visualização suportada pelo *Oracle Business Intelligence Discoverer*.

A apresentação dos resultados obtidos foi apresentada em reunião com a Direção de Serviços e seus Coordenadores de equipa, estando as principais conclusões dessa reunião sintetizadas no final do capítulo, com a exemplificação das potencialidades de extração, análise e exploração de dados.

7.1 *Oracle Business Intelligence Discoverer*

O *Oracle Business Intelligence Discoverer* é uma ferramenta que possibilita a construção intuitiva de interrogações *ad hoc*, a fabricação de relatórios e a análise de dados, tirando partido das características intrínsecas do modelo dimensional de um armazém de dados, como sejam por exemplo, os factos, as dimensões ou as hierarquias. Possui ainda funcionalidades que lhe possibilitam a publicação em linha.

Estas ferramentas proporcionam o acesso imediato à informação contida em armazéns de dados, por pessoas sem conhecimentos na área.

7.1.1 *Arquitetura do Oracle Discoverer (2009)*

A Figura 7.1.1 apresenta uma visão geral da arquitetura da ferramenta, na qual é visível a sua organização em camadas.

Na camada da base de dados, em baixo, encontram-se os dados e metadados relacionados com a *End User Layer* (um conceito explicado na secção 7.1.2) e com as definições dos cadernos de trabalho (um conceito explicado na secção 7.1.3).

A camada intermédia é a camada aplicacional, onde a ferramenta se encontra instalada.

A camada cliente permite o acesso via navegadores de *internet* aos conteúdos disponibilizados pelo *Discoverer*, usando para o efeito duas ferramentas Web, o *Discoverer Plus*, para a edição de interrogações e relatórios e o *Discoverer Viewer*, vocacionada para a análise e exploração de dados.

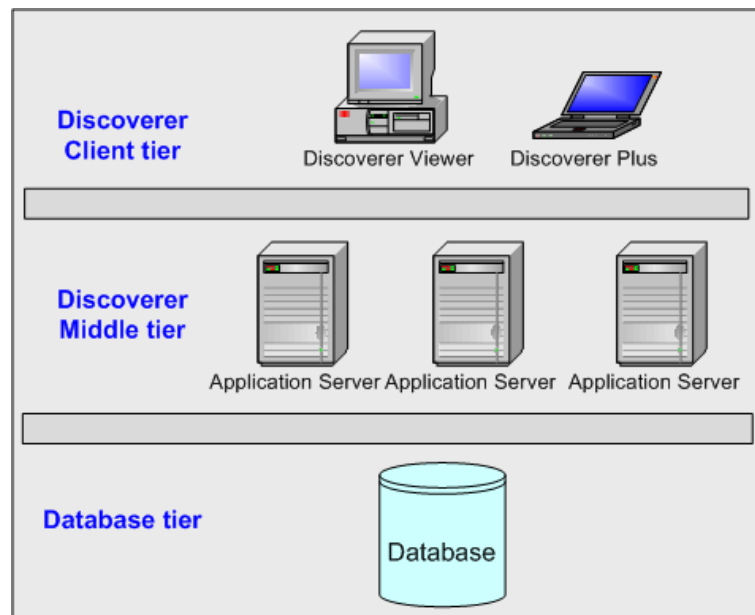


Figura 7.1.1: Arquitetura do *Oracle BI Discoverer* (ORACLE 2009)

7.1.2 *End User Layer*

O primeiro passo para utilizar o *Discoverer* inicia-se com a criação de uma *End User Layer* (EUL). Essa tarefa é conseguida através da ferramenta de administração chamada *Discoverer Administrator*.

A EUL é a componente que fornece uma camada de abstração, possibilitando a construção de interrogações, sem conhecer SQL e sem conhecer a complexidade da estrutura dos dados.

A Figura 7.1.2, abaixo apresentada, pretende ilustrar o objetivo da EUL, mascarar os dados e apresentá-los de forma amigável ao utilizador final.

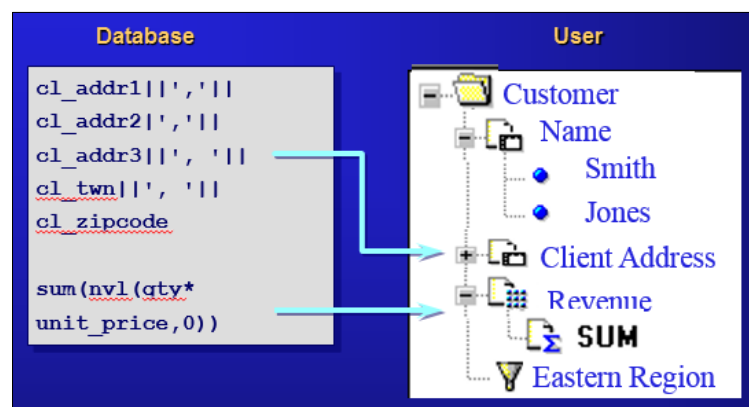


Figura 7.1.2: *End User Layer* (Rose 1995)

7.1.3 Terminologia

O *Discoverer* baseia-se numa série de conceitos que convém conhecer. Esses conceitos estão intimamente relacionados com a forma como os dados são internamente organizados, possuindo uma hierarquia que pretende refletir a estrutura de uma organização.

Assim, o primeiro conceito apresentado é o conceito de área de negócio (*business area*). Uma área de negócio é um agrupamento lógico de tabelas ou vistas da base de dados que satisfazem as necessidades de determinada área.

Essas áreas de negócio são criadas através do *Discoverer Administrator*, agrupando os dados mais adequados, para que seja possível a análise, suporte à decisão e a apresentação de resultados.

No caso específico do armazém de dados implementado para o SDI, as áreas de negócio correspondem às áreas de cada equipa. Desta forma, foram criadas duas áreas de negócio, ERI e EAD, que são as equipas para as quais foram já desenvolvidos *data marts*.

A Figura 7.1.3 apresenta as áreas de negócio criadas.

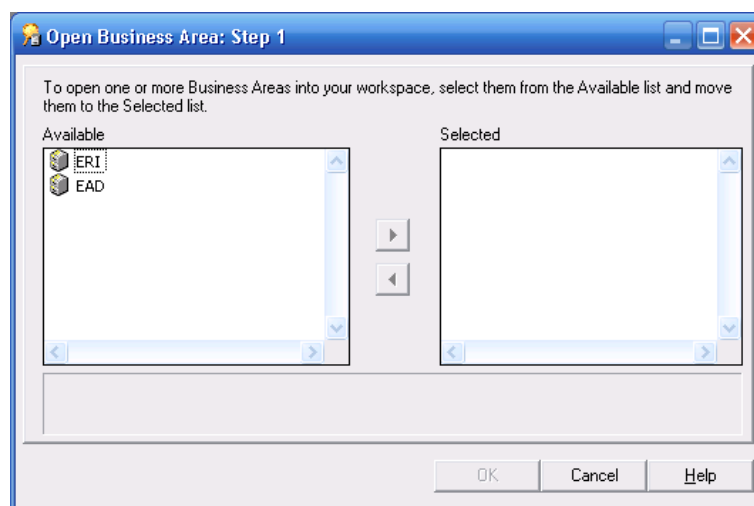


Figura 7.1.3: Áreas de negócio criadas para a apresentação do armazém de dados do SDI

As áreas de negócio são organizadas por pastas (*folders*). As pastas armazenam detalhes sobre grupos de informação relacionada. No *Discoverer*, cada tabela, vista ou combinação das duas corresponde a uma pasta (Rose 1995).

As pastas são compostas por itens, correspondendo cada item a uma coluna de uma tabela da base de dados. Cada item tem um nome e contém um tipo específico de informação.

Depois existem as interrogações, que são formadas por conjuntos de itens e que recuperam informações da base de dados.

A Figura 7.1.4 reflete a estrutura hierárquica acima descrita, configurada para a área de negócio da ERI. Nesta pode ser observada a área de negócio, seguida das pastas e itens ao dispor, para a interrogação da base de dados.

Finalmente aparecem as folhas de trabalho (*worksheets*) que são geradas quando uma interrogação é executada. Contêm a informação recuperada da base de dados, baseada nos itens da interrogação. No *Discoverer* essas folhas de trabalho podem ser armazenadas em cadernos (*workbooks*), facilitando desta forma a organização da informação recuperada. Neste caso concreto, a regra seguida foi a criação de um caderno para cada estrela, acomodando várias folhas de trabalho que tentam responder aos diversos indicadores pretendidos.

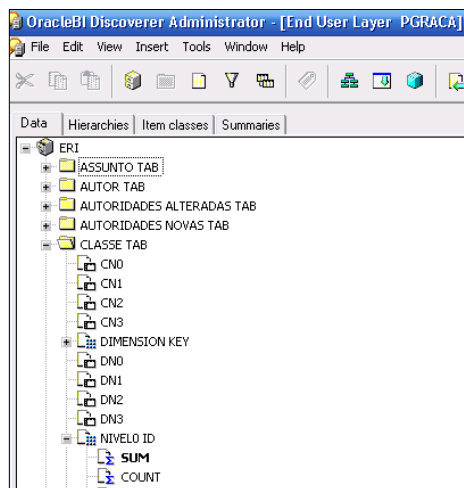


Figura 7.1.4: Organização hierárquica do *Discoverer*

Relativamente às folhas de trabalho, estas podem ainda ser classificadas de acordo com o estilo de apresentação de resultados. O *Discoverer* suporta folhas de trabalho com formato tabela (com colunas e cabeçalhos), folhas de trabalho com formato de referência cruzada (matriz) e ainda folhas de trabalho com o formato de detalhe de página (*page-detail*), formato onde uma tabela ou matriz é produzida para cada um dos itens de página (Rose 1995).

7.1.4 Assistente de criação de folhas de trabalho

A criação das folhas de trabalho que proporcionam a exploração do armazém de dados é realizada no *Discoverer Plus*, a ferramenta de edição e especificação de interrogações e relatórios.

Essa tarefa é realizada recorrendo ao auxílio de um assistente, que permite proceder à definição inicial das mais variadas propriedades da folha, assim como definir a interrogação que está na base do seu relatório.

Do leque das propriedades de possível configuração fazem parte os objetos a apresentar na folha (título, itens de página, área de notas, estilo de apresentação de resultados (matriz ou tabela), presença de gráfico e respetiva disposição na página), tal como pode ser observado na Figura 7.1.5.

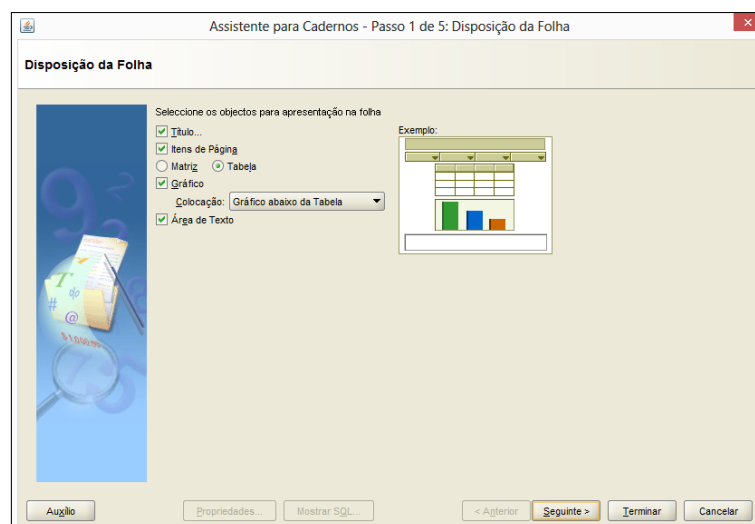


Figura 7.1.5: Seleção de objetos a apresentar numa folha de apresentação de resultados

Os itens, condições, cálculos, percentagens e totais disponíveis para utilização na folha são outro tipo de objetos de configuração obrigatória. Na Figura 7.1.6 pode ser observada a configuração dos itens a utilizar numa folha. Na Figura 7.1.7 pode ser observada a especificação de uma condição que fará parte de uma folha de dados. Mais especificamente, trata-se da condição aplicada à folha de trabalho “Multas” que permite filtrar as transações pelo seu estado (pagas, não pagas ou perdoadas).

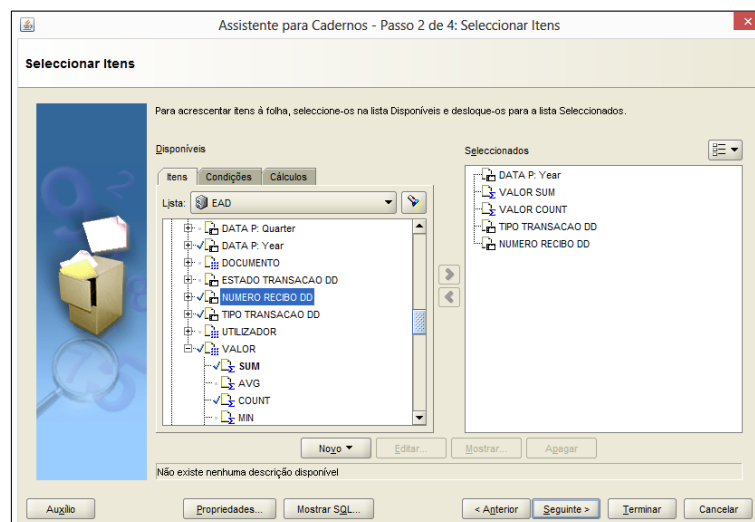


Figura 7.1.6: Seleção dos itens a incluir numa folha de apresentação de resultados

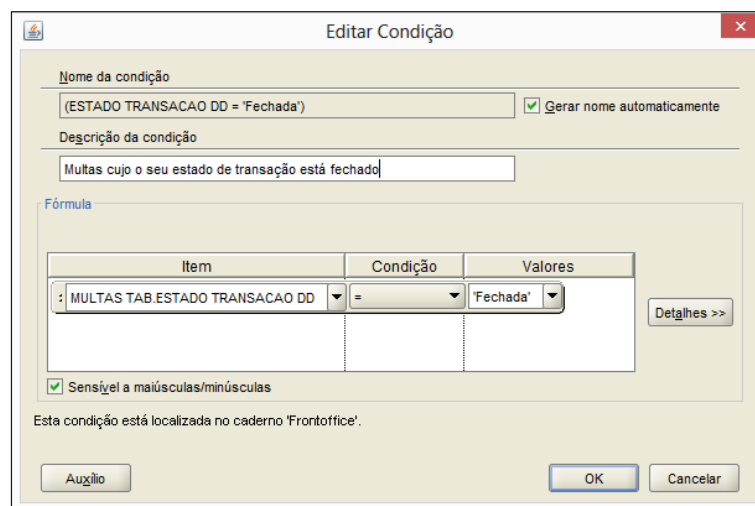


Figura 7.1.7: Especificação de uma condição a incluir numa folha de apresentação de resultados

Segue-se a fase de definição da organização dos itens que compõem a matriz ou tabela da folha, através da alteração da sua disposição, recorrendo à rotação destes entre a área de itens de página, ponto de dados e as linha ou colunas da matriz/tabela em causa. A Figura 7.1.8, abaixo apresentada, demonstra um exemplo da configuração da organização a ser adotada numa folha de trabalho.

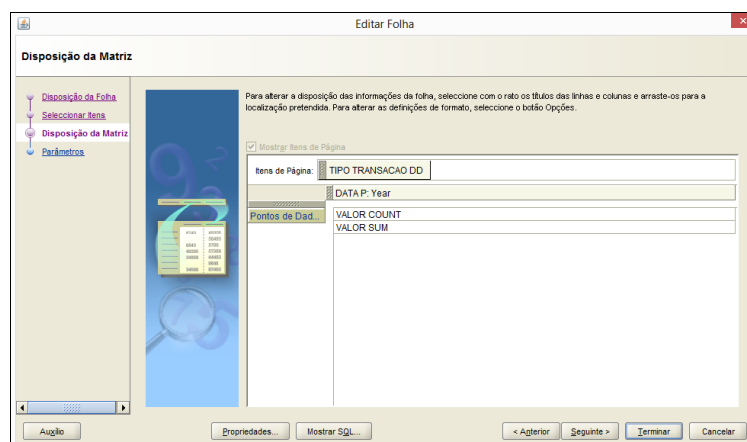


Figura 7.1.8: Especificação da organização de itens numa folha de apresentação de resultados

A criação de parâmetros de entrada que permitam solicitar aos utilizadores a seleção ou introdução de um valor é outro aspeto de configuração possível. A Figura 7.1.9 apresenta um exemplo da definição de um parâmetro para a introdução inicial do ano a partir do qual é baseada a consulta à base de dados.

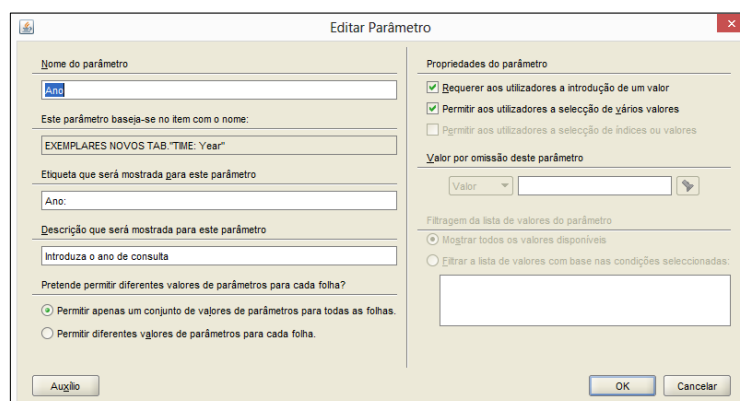


Figura 7.1.9: Especificação de parâmetros de entrada para folhas de apresentação de resultados

Caso o estilo de folha escolhido seja o estilo tabela, é ainda possível configurar ordenações de dados. A Figura 7.1.10 é disso exemplo, apresentando a configuração de uma ordenação para uma folha de trabalho com este estilo.

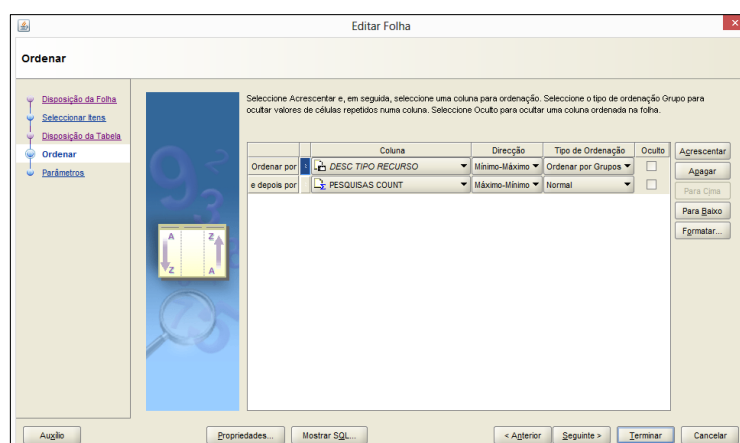


Figura 7.1.10: Especificação de ordenações para uma tabela

Salienta-se que no decorrer do processo de criação de uma folha de trabalho é sempre possível verificar a interrogação SQL resultante. A Figura 7.1.11 mostra o exemplo de uma interrogação SQL, resultante do processo de construção da folha de apresentação de resultados para a folha de trabalho “Circulação”.

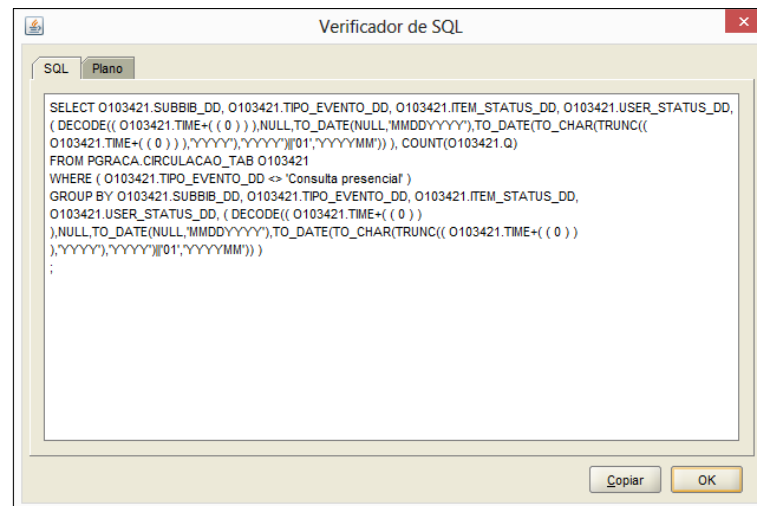


Figura 7.1.11: Interrogação SQL resultante do processo de criação de uma folha de apresentação de resultados

O processo de criação de uma folha de apresentação de resultados no *Discoverer* é um processo iterativo, no qual uma interrogação é criada com a ajuda do assistente, os seus resultados são apresentados, as configurações relacionadas com a apresentação são ajustadas e se no fim o resultado não for o desejado, a interrogação inicial pode ser refinada recorrendo novamente ao assistente de criação de interrogações.

A Figura 7.1.12 exemplifica o resultado final do processo de criação de uma folha de resultados no *Discoverer Plus*, no qual pode ser observado o título da folha na parte superior, a matriz à esquerda e o gráfico resultante à direita.

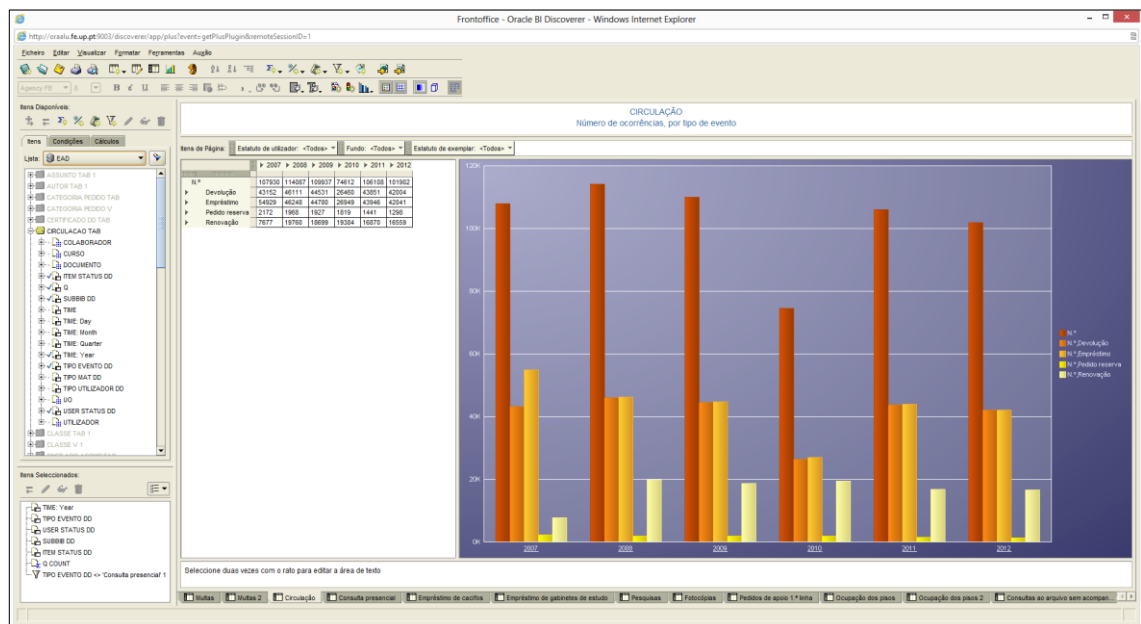


Figura 7.1.12: Folha de apresentação de resultados no *Discoverer Plus*

7.2 Análise e exploração de dados

O armazém até aqui construído possui um elevado volume de dados. Tal quantidade deverá ser manipulada para que a sua apresentação seja perceptível e isenta, tanto quanto possível, de ambiguidades.

7.2.1 Organização e acesso a dados

O processo de criação de folhas de resultados seguiu princípios que tornaram a sua organização coincidente com a estrutura organizativa do Serviço, fazendo corresponder a criação de cadernos do *Discoverer* com os processos detetados na fase de levantamento de requisitos e fazendo corresponder cada estrela do armazém de dados a pelo menos uma folha de trabalho. Desta forma, pretendeu-se que os potenciais utilizadores naturalmente se identificassem com as suas áreas de negócio.

Assim, para as estrelas desenvolvidas, foram criados três cadernos *Discoverer* coincidentes com os processos tratados até ao momento, de acordo com as importâncias e prioridades de implementação apuradas em 5.2.

A Figura 7.2.1, abaixo apresentada, mostra o menu das listas de folhas ao dispor dos utilizadores, resumindo de forma eficiente as folhas de trabalho criadas e respetiva organização.

Foco	Nome	Descrição	Proprietário	Última Modificação
▼	Cadernos do Discoverer			
+	▼ Frontoffice		PGRACA	terça-feira, 11 de Junho de 2013 01:26:49 WEST
	Multas			
	Multas 2			
	Circulação			
	Consulta presencial			
	Empréstimo de cacifos			
	Empréstimo de gabinetes de estudo			
	Pesquisas			
	Fotocópias			
	Pedidos de apoio 1.ª linha			
	Ocupação dos pisos			
	Ocupação dos pisos 2			
	Consultas ao arquivo sem acompanhamento			
	Reserva de gabinetes			
+	▼ Gestão Documental		PGRACA	terça-feira, 11 de Junho de 2013 00:36:04 WEST
	Documentos novos			
	Documentos Alterados			
	Exemplares novos			
	Exemplares alterados			
	Tratamento de documentos			
	Receção Material para tratamento Técnico			
+	▼ Infoteracia		PGRACA	segunda-feira, 10 de Junho de 2013 19:02:31 WEST
	Pedidos de apoio 2.ª linha			
	Pesquisas			
	Lista de recursos pesquisados			
	Lista de recursos pesquisados 2			
Renovar				

Figura 7.2.1: Cadernos e folhas de trabalho desenvolvidas

Para aceder aos dados é usado o *Discoverer Viewer*, uma componente web do *Oracle BI*, orientada para o utilizador final e que tem como objetivo a apresentação, análise e exploração de dados.

7.2.2 Apresentação de dados

Depois de devidamente validado no *Discoverer Viewer*, o utilizador ganha acesso à lista de cadernos e folhas de trabalho, tal como ilustrado na Figura 7.2.1.

Após escolher a opção que pretende explorar, é apresentada ao utilizador a folha de trabalho, previamente configurada no *Discoverer Plus*.

Normalmente a folha de trabalho apresenta uma configuração semelhante à da Figura 7.2.2, com uma área reservada à apresentação do título, uma área reservada à apresentação dos dados sob o formato de uma tabela ou matriz, com eventuais itens de página e com uma área reservada ao gráfico que resulta dos dados apresentados. Neste caso específico, a folha de trabalho mostra os dados relacionados com o registo da receção de material para tratamento técnico, referente a um período de três anos, distribuído pelo seu tipo de material.

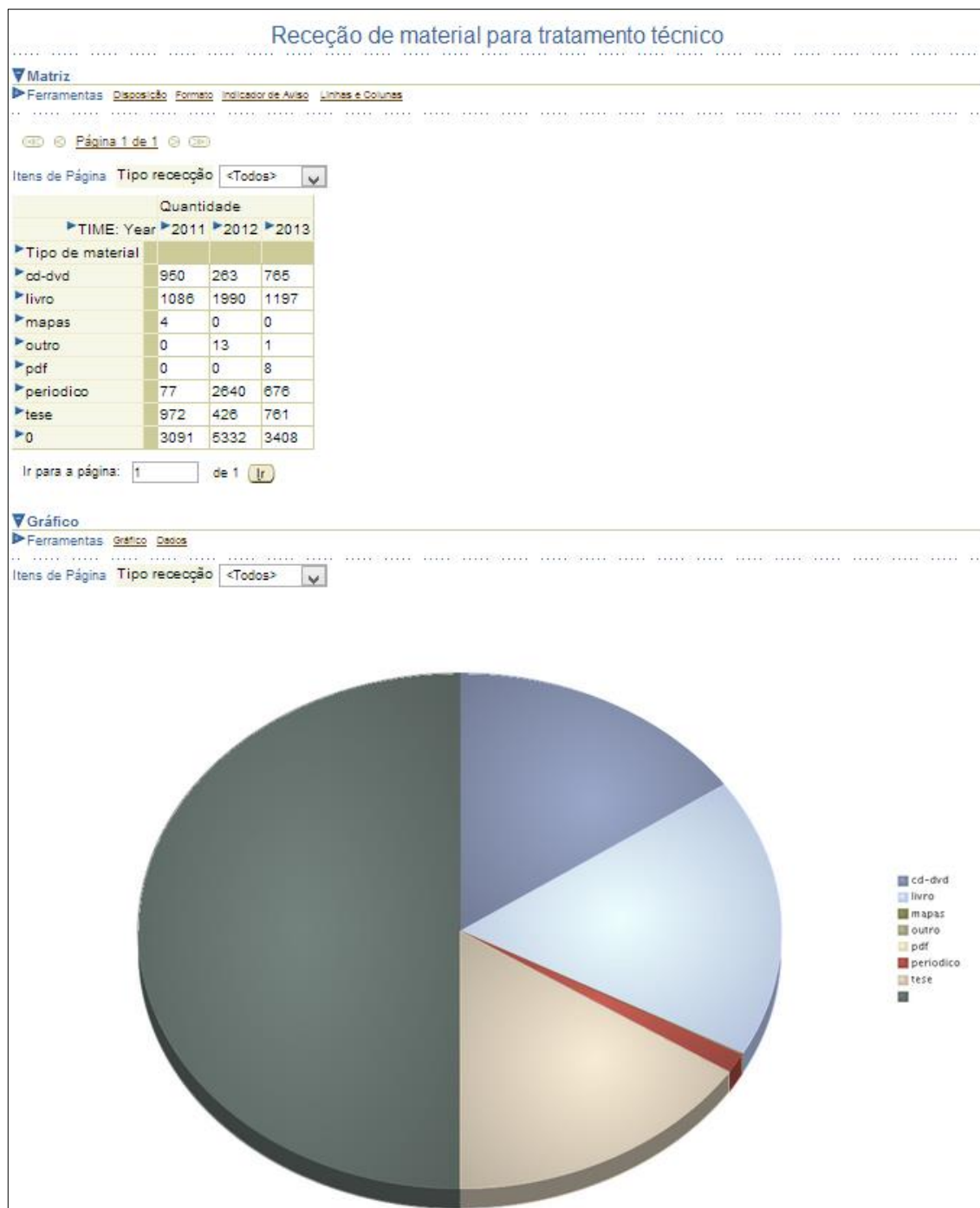


Figura 7.2.2: Apresentação de uma folha de trabalho no *Discoverer Viewer*

Para além da composição previamente determinada é sempre conferido ao utilizador algum grau de liberdade que lhe permite alterar alguns aspetos relacionados com a apresentação dos dados.

Relativamente à matriz ou tabela, o *Discoverer Viewer* permite controlar a disposição das linhas ou colunas, a criação de formatos condicionais, a utilização de indicadores de aviso e o número de linhas e colunas visualizadas.

Apresentam-se em seguida dois exemplos de funcionalidades ao dispor dos utilizadores, úteis para a análise e exploração de dados. O primeiro exemplo, na Figura 7.2.3, é a utilização de formatos condicionais para evidenciar certos dados que correspondam a determinados critérios. No exemplo, o formato configurado evidencia os dados cuja quantidade é inferior a 10, recorrendo à sinalização das células com o fundo vermelho e a fonte a negrito e itálico.

	Quantidade		
▶TIME: Year	▶2011	▶2012	▶2013
▶Tipo de material			
▶cd-dvd	950	263	765
▶livro	1086	1990	1197
▶mapas	4	0	0
▶outro	0	13	1
▶pdf	0	0	8
▶periodico	77	2640	676
▶tese	972	426	761
▶0	3091	5332	3408

Figura 7.2.3: Utilização de formatos condicionais para análise e exploração de dados

O segundo exemplo demonstra a utilização de indicadores de aviso. Um indicador de aviso é um mecanismo semelhante ao formato condicional, mas neste caso é possível definir três intervalos de valores, um superior, um intermedio e outro inferior. O resultado obtido é o realce das células do quadro de resultados recorrendo a um código de cores (verde-amarelo-vermelho). A Figura 7.2.4 é disso exemplo, apresentando o número de pedidos de apoio de 2.^a linha registados no ano de 2013 pela Equipa de Apoio e Descoberta (EAD), distribuído pelo seu tipo de contato.

	QUANTIDADE				
▶Mês	▶Jan	▶Fev	▶Mar	▶Abr	▶Mai
▶Tipo contato					
▶chamada	4	1	4	1	4
▶email	21	24	17	25	10
▶presencial	7	18	21	16	6
▶0	0	0	1	0	1

Figura 7.2.4: Utilização de indicadores de aviso para análise e exploração de dados

A área reservada à exibição do gráfico oferece igualmente algumas possibilidades de controlo. Essas possibilidades assentam essencialmente na escolha do tipo de gráfico utilizado na representação dos resultados e ainda na definição da sua dimensão. Na Figura 7.2.5, abaixo apresentada, é visível a barra de ferramentas, onde é viável a alteração do tipo de gráfico escolhido para a representação de dados. De salientar ainda a tipologia diversificada ao dispor do utilizador.

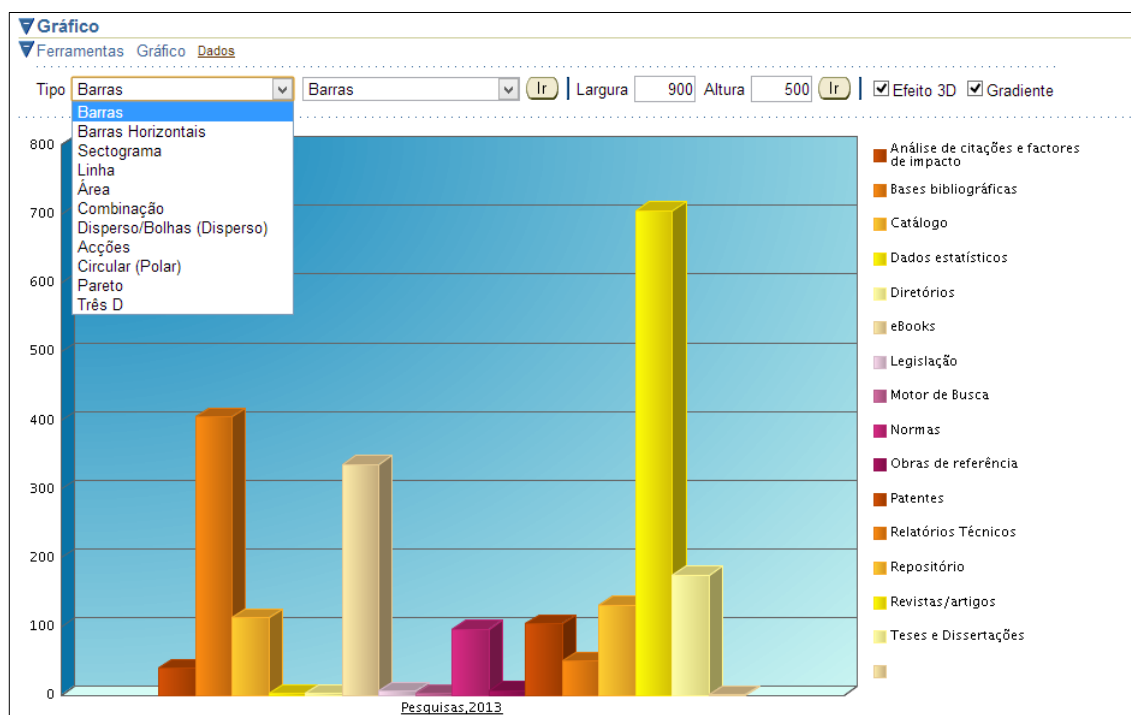


Figura 7.2.5: Alteração do tipo de gráfico durante o processo de análise e exploração de dados

Conclui-se a secção dedicada à apresentação de resultados com a demonstração daquela que se considera ser a mais importante característica em termos de análise e exploração de dados do *Discoverer*. Essa característica tira partido das características do modelo dimensional subjacente, permitindo a navegação pelos dados, fazendo variar a sua granularidade, ao longo da hierarquia de uma dimensão. Esse comportamento é vulgarmente conhecido por *drill up* ou *drill down*.

A navegação nos dados, anteriormente referida, pode ser realizada não só no quadro de resultados, como também no próprio gráfico, sendo que as alterações a um quadro de resultados afetam o gráfico associado e vice-versa.

A Figura 7.2.6, apresentada em baixo, exemplifica a navegação executada na matriz apresentada na Figura 7.2.3. No exemplo apresentado a opção escolhida foi detalhar os resultados de 2012, por trimestre. O resultado pode ser consultado na Figura 7.2.7, onde é visível a expansão nas colunas, sofrida pela matriz e pelo gráfico.

	Quantidade		
	2011	2012	2013
TIME: Year			
TIME: Year			
TIME: Quarter			
TIME: Month			
TIME: Day			
TIME			
Todas as Hierarquias			
Definir Nível de Detalhe do Item Relacionado			
► Tipo de material			
► cd-dvd	950	2	
► livro	1086	1	
► mapas	4	0	
► outro	0	1	
► pdf	0	0	
► periodico	77	2640	676
► tese	972	426	761
► 0	3091	5332	3408

Figura 7.2.6: Navegação de dados(*drill down*)

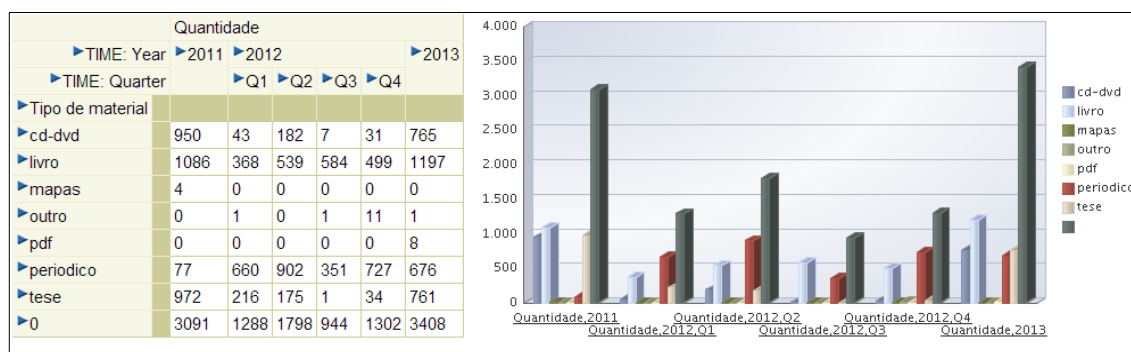


Figura 7.2.7: Resultado da navegação de dados *drill down*

No fim da sua análise, o utilizador dispõe de várias opções, das quais se destacam o envio de resultados por correio eletrónico e a exportação de dados nos mais variados formatos (pdf, xls, csv, png, jpg, xml, etc).

7.3 Reunião de apresentação e validação de resultados

A reunião de apresentação de resultados contou com a participação da Direção de Serviços do SDI e com os Coordenadores das Equipas.

Na ordem de trabalhos da reunião figuravam dois grandes objetivos:

- O primeiro objetivo foi a apresentação da versão atualizada da matriz de barramento dimensional e a versão atualizada da matriz de processos e fontes de dados, procurando desta forma concluir a validação do trabalho de levantamento e projeto desenvolvidos.
- O segundo e principal objetivo foi o de apresentar o trabalho de desenvolvimento realizado até ao momento, através da demonstração de resultados e das funcionalidades da plataforma.

Registou-se uma satisfação generalizada quanto aos resultados obtidos, confirmando-se deste modo a validade e relevância do protótipo desenvolvido, de acordo com os objetivos gerais e específicos delineados para a presente dissertação.

Foi dada conta de casos pontuais relacionados com a diferença entre os dados encontrados e os dados inicialmente esperados e respetivas implicações no modelo idealizado.

Ainda que saindo do âmbito dos objetivos especificados para o desenvolvimento do armazém de dados, foram registadas algumas sugestões pontuais, relacionadas com usabilidade, por parte dos elementos integrantes da reunião, relativamente ao interface de consulta de dados, nomeadamente a sua simplificação.

7.4 Conclusões

A apresentação de resultados, os do armazém de dados e por consequência os do trabalho realizado, é o culminar de um trajeto já longo e que finalmente se manifesta de forma perceptível.

Este capítulo tratou o tema da apresentação de resultados, iniciando com a exposição da ferramenta eleita para interagir com os utilizadores, dando especial enfoque à demonstração das suas funcionalidades. Por fim, o capítulo finaliza com a apresentação das conclusões, resultantes da reunião realizada para a apresentação do protótipo desenvolvido e que culminou com a sua aprovação.

8 Conclusão

A presente dissertação pretendeu dar resposta ao problema inicialmente formulado, que consistia em dotar o SDI de ferramentas e mecanismos necessários para suportarem o processo de tomada de decisão de gestão de forma eficaz.

A revisão da bibliografia evidenciou carências que confirmaram a formulação do problema, motivação e objetivos que estão na génese desta dissertação e que se traduzem por:

- Uma ausência de discussão quanto ao modelo a adotar para um armazém de dados, coerente com um serviço de documentação e informação universitário;
- Soluções unicamente focadas no domínio das bibliotecas;
- Ausência de debate para caracterização dos serviços complementares (museu, editorial, arquivo), no que diz respeito às origens de dados e a indicadores de medida que os caracterizem.

A resposta às carências evidenciadas passaram por:

- A formulação de uma proposta de um modelo dimensional, apropriado ao desenvolvimento de um armazém de dados, que se constitua como uma ferramenta útil para o apoio à decisão de um serviço de documentação e informação universitário;
- A implementação de um armazém de dados para o SDI da FEUP, para a extração de indicadores de medida fiáveis e de qualidade.

Foi utilizada uma metodologia própria de um projeto de engenharia informática, através da qual as respostas aos objetivos gerais e específicos inicialmente formulados foram sendo progressivamente dadas.

A realização do estudo orgânico-funcional ao SDI possibilitou o levantamento dos processos e das fontes de dados que lhes estavam associadas, materializado numa matriz que resume a arquitetura completa de processos do Serviço, na perspetiva das equipas e nos dados que os suportam. A matriz foi construída em colaboração com o SDI, validada pelos coordenadores das equipas e os processos classificados quanto à sua importância e prioridade de implementação. Consideramos esta matriz um primeiro resultado desta dissertação.

Foi também realizado um levantamento dos indicadores de medida considerados relevantes pelos coordenadores, como um passo no sentido da definição das medidas a incluir nas estrelas do armazém de dados.

Com base na informação compilada, foi desenvolvido o modelo dimensional para a implementação do armazém de dados. Esse modelo é apresentado sob a forma de uma matriz de barramento dimensional, acompanhada pelos respetivos dicionários de factos e dimensões. Esta

segunda matriz tem também um caráter de cobertura completa do Serviço e constitui um segundo resultado da dissertação, que pode ser útil em projetos similares noutros Serviços de Documentação e Informação universitários.

Atendendo à extensão do problema e aos recursos disponíveis, a fase de desenho mais detalhado cingiu-se aos *data marts* considerados prioritários. Foram identificadas as estrelas mais relevantes que visam dar resposta aos indicadores selecionados. Não existindo na literatura nenhum modelo adequado ao serviço, não foi possível efetuar a comparação do trabalho desenvolvido com outras realidades. Este modelo constitui o terceiro resultado da dissertação.

O quarto resultado é um protótipo que implementa duas dezenas de estrelas, correspondendo a dois *data marts*, com dados entre 2007 e 2012. Sobre estes dados foi elaborada uma camada de apresentação para a visualização tabular e gráfica de alguns dos indicadores definidos. O protótipo implementado seguiu os modelos idealizados, assentando nas tecnologias inicialmente estudadas e propostas. Demonstra que apesar de não existir documentação de outras implementações é adequado o seu desenvolvimento no contexto dos serviços de documentação e informação universitários.

O processo de conceção do armazém de dados contou com o estreito e regular contributo dos coordenadores de equipas do SDI, de acordo com a metodologia proposta, através da realização frequente de reuniões, onde o trabalho desenvolvido foi sendo discutido e reavaliado.

O armazém de dados implementado permite obter os indicadores e contribuir para um processo de decisão mais informado. Responde-se assim ao problema inicialmente colocado, justificando o trabalho desenvolvido e motivando o seu aprofundamento futuro.

Durante o processo de implementação, foram encontrados diversos obstáculos, ora de natureza técnica, ora relacionados com os dados disponíveis. Os obstáculos de natureza técnica foram ultrapassados recorrendo às soluções alternativas encontradas. Os obstáculos relacionados com os dados prendem-se essencialmente com a disparidade entre o modelo idealizado e a realidade dos dados existentes. Nos casos em que essa diferença foi detetada, o modelo inicialmente projetado foi ajustado de forma a oferecer soluções viáveis.

Como forma de superar algumas dessas dificuldades, assegurando a qualidade indispensável dos dados, recomenda-se que o SDI explore ao máximo os sistemas operacionais existentes em detrimento do registo manual de informação.

A apresentação dos resultados, realizada no âmbito de uma reunião com a Direção de Serviços e os coordenadores das equipas, validou o trabalho desenvolvido, registando-se a aceitação unânime do trabalho efetuado, sendo desta forma confirmada a exequibilidade e utilidade do armazém de dados.

TRABALHO FUTURO

O desenvolvimento realizado com o armazém de dados até ao momento centrou-se na implementação de um protótipo para dar resposta aos objetivos da dissertação.

O aprofundamento do seu desenvolvimento traz necessariamente desafios e oportunidades que contribuirão para o enriquecimento de um trabalho desta natureza.

Deste modo, deve ser equacionada numa primeira fase, a migração da infraestrutura desenvolvida, do servidor de base dados destinado a trabalhos de alunos para um servidor dedicado, reavaliando se necessário o suporte tecnológico a adotar, se o fator financeiro for determinante.

Devem ser investidos esforços no desenvolvimento de mecanismos que possibilitem uma eficiente recolha e atualização de dados.

Deve ser dado seguimento ao processo de aprofundamento do modelo dimensional que está na base do armazém de dados, tendo como eixos principais de estudo as suas dimensões e os seus factos.

A interface com o utilizador deverá ser também objeto de aprofundamento, dado que os trabalhos realizados nesta dissertação apenas se circunscreveram ao cumprimento dos seus

objetivos. O investimento na interface com o utilizador pode inclusive originar trabalhos na área de testes e usabilidade.

Durante a apresentação do protótipo, à medida que os utilizadores interagem com os dados apresentados, registou-se a possibilidade de existirem ainda outros indicadores de medida por descobrir e que só serão revelados com o contacto dos utilizadores com o sistema. O desenvolvimento de novos indicadores não pode ser descartado.

Finalmente, este armazém de dados pode ser encarado como a pedra basilar para novas oportunidades de investigação, com estudos nas áreas do desenvolvimento de aplicações de sistemas de suporte a decisão ou ainda na área da inteligência artificial, onde a presença de armazéns de dados é frequente.

Referências bibliográficas

- AGORA. 2009. Regulamento Identitário e Operacional. Porto: FEUP.
- Allan, David. 2010. "OWB 11gr2 – Degenerate Dimensions". Acedido a 8 de Junho de 2013. https://blogs.oracle.com/warehousebuilder/entry/owb_11gr2_degenerate_dimensions.
- Awan, A. e F. Daneshgar. 2011. "A methodology for development of a model warehouse for identifying innovative services in academic libraries". Comunicação apresentada em Information Science and Service Science (NISS), 2011 5th International Conference on New Trends in. 24-26 Oct. 2011.
- Bartolini, Gabriele. 2009. "Data warehousing with PostgreSQL". Comunicação apresentada em European PostgreSQL Day, em ParisTech Telecom, Paris, France.
- Berkus, Josh. 2011. "Data Warehousing with PostgreSQL". Comunicação apresentada em MySQL User Conference. 2011.
- Costa, Fábio Pinto da e Pedro Graça. 2012. Desenho e Implementação de um armazém de dados. Faculdade de Engenharia da Universidade do Porto.
- Decker, Reinhold e Michael Höppner. 2006. "Strategic planning and customer intelligence in academic libraries." *Library Hi Tech* no. 24 (4):504 - 514. doi: 10.1108/07378830610715374
- 10.1108/14635781111150376
- 10.1108/14635781111150286
- 10.1108/14635781111150312.
- Di Giacomo, M. 2005. "MySQL: Lessons learned on a digital library." *IEEE Software* no. 22 (3):10-13. <http://www.scopus.com/inward/record.url?eid=2-s2.0-21244443476&partnerID=40&md5=7d5bd447c765b1f5f85940286ecb57b4>. doi: 10.1109/ms.2005.71.
- EAD. 2009. Qual a nossa identidade? Como funcionamos? Porto: FEUP.
- EMAP. 2009. Equipa da Memória e Acesso Perene: Modelo identitário e funcional. Porto: FEUP.
- EnterpriseDB. 2010. Postgres Plus 8.4 vs. MySQL 5.5. Apresentação, Em *Feature Comparison and Commentary*. <http://www.enterprisedb.com>.
- ERI. 2009. Regulamento. Porto: FEUP.
- ESA. 2009. Regulamento Identitário e Operacional. Porto: FEUP.
- ESE. 2009. Documento Identitário. Porto: FEUP.
- ExLibris Group. 2007. Aleph Brochure. Brochure, editado por ExLibris. http://www.exlibrisgroup.com/files/Products/Aleph.Voyager/Aleph/Aleph_A4_low.pdf.

- . 2012. ARC User's Guide. User Manual: ExLibris. [http://customercenter.exlibrisgroup.com/DocumentationCenter/Ex%20Libris%20Docu mentation/Aleph/Technical%20Documentation/Aleph%20Reporting%20Center/Versio n%203.0/Staff%20User%20Information/ARC%203.0%20User%20Guide.pdf](http://customercenter.exlibrisgroup.com/DocumentationCenter/Ex%20Libris%20Documentation/Aleph/Technical%20Documentation/Aleph%20Reporting%20Center/Versio n%203.0/Staff%20User%20Information/ARC%203.0%20User%20Guide.pdf).
- Faculdade de Engenharia da Universidade do Porto. "Serviço de Documentação e Informação". Acedido a 22 de Abril de 2012. http://sigarra.up.pt/feup/unidades_geral.visualizar?p_unidade=84.
- . 2004. Regulamento orgânico dos serviços centrais da Faculdade de Engenharia da Universidade do Porto. editado por Reitoria da Universidade do Porto. Lisboa: Diário da Republica.
- FEUP Edições. 2009a. Modelo do Processo Comercializar Obra. Porto: FEUP.
- . 2009b. Modelo do Processo Publicar Obra. Porto: FEUP.
- IBM. 2012. "IBM - Cognos Business Intelligence - Software". Acedido a 7 de Julho de 2012. <http://www-01.ibm.com/software/analytics/cognos/business-intelligence/>.
- IPQ. 2004. *Norma Portuguesa ISO 11620*. Instituto Português da Qualidade.
- . 2006. *Norma Portuguesa ISO 11620:2004/A1*. Instituto Português da Qualidade.
- ISO. 2003. *ISO2789 - Information and documentation - International library statistics*. Switzerland: International Organization for Standardization. Acedido a 12 de Julho de 2012.
- Kimball, Ralph. 1998. *The Data Warehouse lifecycle toolkit expert methods for designing, developing, and deploying Data Warehouses*. New York [etc.]: John Wiley & Sons.
- . 2011. "Design Tip #136 Adding a Mini-Dimension to a Bridge Table". Acedido a 8 de Junho de 2013. <http://www.kimballgroup.com/2011/06/28/design-tip-136-adding-a-mini-dimension-to-a-bridge-table/>.
- Kimball, Ralph e Margy Ross. 2002. *The data warehouse toolkit the complete guide to dimensional modeling*. Vol. 2nd ed. New York [etc.]: John Wiley & Sons.
- Lumpkin, George. 2009. Oracle Database 11g for Data Warehousing and Business Intelligence White Paper. White paper: Oracle Corporation.
- Microsoft Corporation. 2012a. "Data Warehousing | Microsoft SQL Server 2008 R2". Acedido a 7 de Julho de 2012. <http://www.microsoft.com/sqlserver/en/us/solutions-technologies/data-warehousing.aspx>.
- . 2012b. SQL Server 2008 R2 Data Warehouse Data Sheet. <http://www.microsoft.com/sqlserver/en/us/solutions-technologies/data-warehousing.aspx>.
- MySQL. 2011a. MySQL datasheet. editado por Oracle. <http://www.mysql.com/products/enterprise/>.
- . 2011b. MySQL Enterprise Edition White Paper. editado por Oracle Corporation and/or its affiliates.
- . 2012a. "MySQL in Data Warehousing & Business Intelligence". Acedido a 8 de Julho de 2012. <http://www.mysql.com/why-mysql/data-warehouse.html>.
- . 2012b. "MySQL Technical Specifications". Acedido a 8 de Julho de 2012. <http://www.mysql.com/products/enterprise/techspec.html>.
- Neves, João Pedro Moura. 2007. "Desenvolvimento de um sistema de apresentação de indicadores de desempenho para serviços de biblioteca ou centros de documentação (INQUIDOC) na Datinform – Informática serviços e estudos S.A.", Faculdade de Letras, Faculdade de Engenharia, Licenciatura de Ciência da Informação, Universidade do Porto.
- Nicholson, S. 2003. "The Bibliomining Process: Data Warehousing and Data Mining for Library Decision-Making." *Information Technology and Libraries* no. 22 (4).
- Nicholson, Scott. "Introducing to the Bibliomining Process". Acedido a 30 de Junho de 2012. http://groups.lis.illinois.edu/guest_lectures/bibliomining/img0.html.
- . 2004. "A conceptual framework for the holistic measurement and cumulative evaluation of library services." *Proceedings of the American Society for Information Science and Technology* no. 41 (1):496-506. <http://dx.doi.org/10.1002/meet.1450410158>. doi: 10.1002/meet.1450410158.

- . 2006. "The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services." *Information Processing & Management* no. 42 (3):785-804. <http://www.sciencedirect.com/science/article/pii/S0306457305000658>. doi: 10.1016/j.ipm.2005.05.008.
- Nicholson, Scott e R. David Lankes. 2007. "The Digital Reference Electronic Warehouse Project." *Reference & User Services Quarterly* no. 46 (3):45-59. <http://search.ebscohost.com/login.aspx?direct=true&db=ehh&AN=24854715&lang=pt-br&site=ehost-live>.
- Nicholson, Scott e Jeffrey M. Stanton. 2003. "Gaining Strategic Advantage through Bibliomining: Data Mining for Management Decisions in Corporate, Special, Digital, and Traditional Libraries." Em.: Hershey, PA: Idea Group Publishing.
- ORACLE. 2001. "Ask Tom "Function to concatenate output"". http://asktom.oracle.com/pls/asktom/f?p=100:11:::P11_QUESTION_ID:229614022562.
- . 2009. *Oracle® Fusion Middleware - Configuration Guide for Oracle Business Intelligence Discoverer 11g Release 1 (11.1.1)*. ORACLE.
- . 2010. "Oracle® Database 2 Day + Data Warehousing Guide 11g Release 2 (11.2)(e10578)." <http://www.oracle.com/technetwork/database/enterprise/11g2day-112010-01.pdf>.
- Pal, Jiban K. 2011. "Usefulness and applications of data mining in extracting information from different perspectives." *Annals of Library and Information Studies* no. 58 (March, 2011):7-16.
- Paynter, Robin A. 2008. "Commercial Library Decision Support Systems: An Analysis Based on Collection Managers' Needs." *Collection Management* no. 34 (1):31-47. Acedido a 23 de Maio de 2012. <http://dx.doi.org/10.1080/01462670802548585>. doi: 10.1080/01462670802548585.
- Pentaho. 2008. Pentaho Data Integration Screenshot. <http://wiki.pentaho.com/display/EAI/Pentaho+Data+Integration+Screenshots>. <http://wiki.pentaho.com/display/EAI/Pentaho+Data+Integration+Screenshots>.
- . 2012. "Pentaho Data Integration". Acedido a 8 de Julho de 2012. <http://www.pentaho.com/explore/pentaho-data-integration/>.
- Pentaho Community. 2012. "Pentaho Community Wiki Home". Acedido a 8 de Julho de 2012. <http://wiki.pentaho.com/display/COM/Community+Wiki+Home>.
- PostgreSQL. WHY USE POSTGRESQL? pgfoundry.org/docman/view.php/1000047/294/lisa_06.pdf.
- Prakash, K, Prem Chand e Umesh Gohel. 2004. "Application of Data Mining in Library and Information Services."
- Project COUNTER. 2012. "COUNTER - Counting Online Usage of Networked Electronic Resources | Home". Acedido a 12 de Julho de 2012. <http://www.projectcounter.org/index.html>.
- Rose, Julie. 1995. Enabling End-User Analytical Capabilities. ORACLE.
- Serial Solutions. 2012a. Data population service for Management Solutions. Bruchura. <http://www.serialssolutions.com/en/resources/detail/data-population-service>.
- . 2012b. "Serial Solutions Product Home Page". Acedido a 4 de Julho de 2012. <http://www.serialssolutions.com/en/solutions/manage>.
- . 2012c. "Ulrich's™ Serials Analysis System Web Page". Acedido a 5 de Julho de 2012. <http://www.serialssolutions.com/en/services/ulrichs/ulrichs-serials-analysis-system>.
- . 2012d. Ulrichs Global Serials Intelligence. <http://www.serialssolutions.com/en/resources/detail/ulrichs-brochure>.
- Shachar-Abutbul, Idit. 2011. ARC Basic Training. Apresentação/Manual de formação, editado por ExLibris Group. <http://customercenter.exlibrisgroup.com>.
- Shieh, Jiann-Cherng. 2010. "The integration system for librarians' bibliomining." *Electronic Library, The* no. 28 (5):709 - 721. doi: 10.1108/02640471011081988 10.1108/14635781111150286

10.1108/14635781111150376

10.1108/14635781111150312.

Swets. 2008. ScholarlyStats Description of Services. editado por Swets: Sweets.
<http://www.swets.com/scholarlystats>.

Thomson Reuters. 2009. "Journal Use Reports™ tutorial". Acedido a 4 de Julho de 2012.
<http://scientific.thomson.com/tutorials/jur/index.html>.

University of Pennsylvania Library. "Penn Library Data Farm Home Page". Acedido a 30 de Junho 2012. <http://datafarm.library.upenn.edu/>.

———. 2012. "metridoc - Data integration tools to assist libraries with collecting business metrics and usage data - Google Project Hosting". Acedido a 30 de Junho de 2012.
<http://code.google.com/p/metridoc/>.

Zucca, Joseph. 2010. "Building Frameworks of Organizational Intelligence".

Anexos

Anexo A. Bases de Dados

A1. Especificações técnicas *MySQL*

Tabela 8: Especificações técnicas do MySQL (MySQL 2012b)

MySQL Technical Specifications		
Flexible Architecture <ul style="list-style-type: none"> • Open Source • Multi-threaded • Pluggable Storage-Engine Architecture • InnoDB • MyISAM • Merge • Memory • Archive • Cluster ANSI SQL Standards <ul style="list-style-type: none"> • ANSI SQL • SubQueries, Joins, Cursors • Prepared Statements • Views • Triggers • Stored Procedures • User-Defined Functions • Information Schema • Unicode • Referential Integrity • Server-enforced Data Integrity Replication & High-Availability <ul style="list-style-type: none"> • Built-in Replication Engine • Master/Slave • Dual Master • Replication Ring • Replication Tree • Row-based Replication 	High-Performance <ul style="list-style-type: none"> • Dynamic Memory Caches • Unique Query Cache • Multiple Index Type (B-tree, R-tree, Hash, etc.) • Cost-based Optimizer • High-speed Query Optimizer • Diagnostics and SQL Tracing • Performance Schema OLTP and Transactions <ul style="list-style-type: none"> • ACID Transactions • Commit, Rollback • Foreign Keys • Referential Integrity • Row-level Locking • Customizable Lock Isolation Levels • Distributed Transactions (XA) • Snapshot • Repeatable Reads (readers don't block writers and vice-versa) • Timed Deadlock Detection Web Application Optimized features <ul style="list-style-type: none"> • Server-side Thread Pool 	Manageability and Ease of Use <ul style="list-style-type: none"> • Easy Install and Setup • "15 minutes to Success" • Information Schema • Event Scheduler MySQL Enterprise Backup¹ <ul style="list-style-type: none"> • Hot Backup for InnoDB • Logical, Physical Backup • Compressed Backup • Full, Incremental, Partial Backups • Full, Partial Restore • Point in Time Recovery • Auto-Restart/Recovery MySQL Enterprise High Availability¹ <ul style="list-style-type: none"> • Oracle VM Template for MySQL • Windows Failover Clustering for MySQL MySQL Enterprise Scalability¹ <ul style="list-style-type: none"> • MySQL Thread Pool MySQL Enterprise Security¹ <ul style="list-style-type: none"> • MySQL External Authentication • PAM (Pluggable Authentication Module) • Windows Native Authentication MySQL Enterprise Monitor¹

MySQL Technical Specifications		
<ul style="list-style-type: none"> • Semi-synchronous Replication • MySQL Cluster HA MySQL Cluster <ul style="list-style-type: none"> • 99.999% Availability • Distributed architecture • Synchronous replication • Real-time transactional performance • SQL & Non-SQL data access • Java, C++, memcached, HTTP Storage Management <ul style="list-style-type: none"> • Storage Administration • Automatic Space Expansion • Built-in Space Reclamation • Auto Undo/Rollback Management • Customizable Storage Assignments • Compressed and Archive Tables • Packed Index Storage • Works with NAS or SAN storage Security <ul style="list-style-type: none"> • SSL Support • Built-in Data Encryption/Decryption • Fine Grained Object Privileges • View Support • Triggers for auditing • Query Logs for auditing 	<ul style="list-style-type: none"> • Connection Thread Caching • Query Results Caching • Load Balancing • Prepared Statements • Full Text Search • Multiple Character Sets • Geospatial Support • Multi-Terabyte Scalability Data Warehouse Optimized features <ul style="list-style-type: none"> • Fast Data Load Utility • High-Speed Multi-Insert Function • GROUP BY WITH ROLLUP • Aggregate UDF • Analytic SQL Functions • Geospatial Support • Multi-Terabyte Scalability Partitioning (Table/Index) <ul style="list-style-type: none"> • Range • Hash • List • Key • Composite 	<ul style="list-style-type: none"> • MySQL Enterprise Dashboard • MySQL Query Analyzer • MySQL Enterprise Advisors • MySQL Replication Monitor MySQL Enterprise Advisors¹ <ul style="list-style-type: none"> • Administration Advisor • Security Advisor • Replication Advisor • Performance Advisor • Memory Usage Advisor • Schema Advisor • Upgrade Advisor • Custom Advisor Graphical Tools <ul style="list-style-type: none"> • MySQL Workbench • Data Modeling • Database Administration • SQL Editor Drivers <ul style="list-style-type: none"> • MySQL Native C Library • MySQL Drivers for ODBC, JDBC, .Net, C, C++ • Community Drivers for PHP, Perl, Python, Ruby Operating Systems <ul style="list-style-type: none"> • Linux, Solaris, FreeBSD, Mac OS, Windows

¹ Características só disponíveis em edições comerciais

A2. Quadro comparativo das características de Bases de Dados

Quadro comparativo das características das quatro bases de dados apresentadas neste estudo (Di Giacomo 2005).

A comparison of features for four database engines				
	MySQL	PostgreSQL	Oracle	MSSQL
Data storage				
Storage models	MyISAM, InnoDB, Berkeley DB, full-text	Postgres	Bitmapped, B-tree, IOT, function-based	Clustered, nonclustered
Reliability	High/very high	High	High/very high	High/very high
Scalability	Large/very large	Large	Large/very large	Large/very large
Indexes				
Single- and multcolumn, primary key, and full text	Yes	Yes	Yes	Yes
Data integrity				
ACID compliance, row-level locking, hot backup, and partial rollback	Yes	Yes	Yes	Yes
Replication				
Single master	Yes	Yes	Yes	Yes
Multimaster	Yes	Yes/no *	Yes	Yes
Clustering	Yes	No	Yes	Yes
Interface methods				
ODBC/JDBC, C/C++, and Java	Yes	Yes	Yes	Yes
Advanced features				
Stored procedures, views, triggers, sequences, and cursors	Yes (starting with version 5.x)	Yes	Yes	Yes
<i>*Solutions exist but they are commercial.</i>				

Anexo B. Indicadores de medida

B.1 Norma NP ISO 11620

Tabela 9: Norma ISO11620 – Grupo B1, indicadores de percepção do utilizador

Grupo	Descrição	Código
Geral	Satisfação do Utilizador	B.1.1.1

Tabela 10: Norma ISO11620 – Grupo B2, indicadores de serviços públicos

Grupo	Descrição	Código
Geral	Percentagem população alvo atingida	B.2.1.1
	Custo por utilizador	B.2.1.2
	Visitas à biblioteca per-capita	B.2.1.3
	Custo por visita à biblioteca	B.2.1.4
Fornecimento de Documentos	Disponibilidade de títulos	B.2.2.1
	Disponibilidade de títulos solicitados	B.2.2.2
	Percentagem de títulos solicitados na coleção	B.2.2.3
	Disponibilidade alargada de títulos solicitados	B.2.2.4
	Consultas na biblioteca per-capita	B.2.2.5
	Taxa de utilização de documentos	B.2.2.6
	Proporção da coleção não-utilizada	B.2.2.7
	Proporção na arrumação de documentos	B.2.2.8
Recuperação de documentos	Tempo médio de recuperação de documentos em depósito	B.2.3.1
	Tempo médio de recuperação de documentos em acesso livre	B.2.3.2
Empréstimo	Movimento da coleção de empréstimo	B.2.4.1
	Empréstimos per-capita	B.2.4.2
	Documentos em empréstimo per-capita	B.2.4.3
	Empréstimos por empregado	B.2.4.4
	Proporção da coleção em empréstimo	B.2.4.5
Fornecimento de documentos de fontes de informação externas	Rapidez do empréstimo interbibliotecas	B.2.5.1
Serviços de referência e formação	Taxas de respostas corretas	B.2.6.1
Pesquisa de informação	Taxa de sucesso da pesquisa no catálogo de títulos	B.2.7.1
	Taxa de sucesso da pesquisa no catálogo de assuntos	B.2.7.2
Formação de utilizadores	Não descrito na norma	
Equipamentos	Disponibilidade dos equipamentos	B.2.9.1
	Taxa de utilização dos equipamentos	B.2.9.2
	Taxa de ocupação de lugares	B.2.9.3
	Disponibilidade dos sistemas automatizados	B.2.9.4

Tabela 11: Norma ISO11620 – Grupo B3, indicadores de serviços técnicos

Grupo	Descrição	Código
Aquisição de documentos	Tempo médio da aquisição de documentos	B.3.1.1
Processamento de documentos	Tempo médio do processamento de documentos	B.3.2.1

Grupo	Descrição	Código
Catálogo	Custos por título catalogado	B.3.3.1

Tabela 12: Norma ISO11620 – Grupo B5, indicadores de disponibilidade e utilização de recursos humanos

Grupo	Descrição	Código
Disponibilidade e utilização de recursos humanos	Pessoal dos serviços ao utilizador per-capita	B.5.1.1
	Percentagem de pessoal dos serviços ao utilizador em relação ao total de empregados	B.5.1.2

B.2 Relatórios COUNTER

Tabela 13: Relatórios COUNTER para Monografias e Obras de Referência

Código	Descrição
BR1	Número de pedidos de títulos bem-sucedidos, por mês e título.
BR2	Número de pedidos de secção bem-sucedidos, por mês e título.
BR3	Número de acessos negados por mês e título.
BR4	Número de acessos negados por mês e serviço.
BR5	Total de pesquisas e sessões por mês e título.
BR6	Total de pesquisas e sessões por mês e serviço

Tabela 14: Relatórios COUNTER para revistas e bases de dados

Código	Descrição
JR1	Número de pedidos de acesso bem-sucedidos a artigos em texto integral, por mês e revista.
JR1a	Número de pedidos de acesso bem-sucedidos a artigos em texto integral de um arquivo, por mês e revista.
JR2	Número de acessos negados por mês e revista.
JR3	Número de pedidos de acesso bem-sucedidos por mês, revista e tipo de página.
JR4	Total de pesquisas por mês e coleção
JR5	Número de pedidos de acesso bem-sucedidos a artigos em texto integral, por ano e revista.
DR1	Total de pesquisas e sessões, por mês e base de dados.
DR2	Número de acessos negados por mês e base de dados.
DR3	Total e pesquisas e sessões, por mês e serviço.
CR1	Número de pedidos de acesso bem-sucedidos a artigos de revista ou capítulos de livros, em texto integral, por mês.
CR2	Total e pesquisas, por mês e base de dados.

Tabela 15: Relatórios COUNTER para Monografias e Trabalhos de Referência

Código	Descrição
BR1	Número de pedidos de acesso bem-sucedidos a Títulos, por mês e título.
BR2	Número de pedidos de acesso bem-sucedidos a secções, por mês e título.
BR3	Número de acessos negados, por mês e título.
BR4	Número de acessos negados, por mês e serviço.
BR5	Número de pesquisas e sessões, por mês e título.
BR6	Número total de pesquisas e sessões, por mês e serviço.

Anexo C. Processos e casos de uso

C.1 EQUIPA DE APOIO E DESCOBERTA (EAD)

Tabela 16: Sumário dos processos/atividades da equipa EAD

Processos	Atividades	Descrição
Frontoffice	<ul style="list-style-type: none"> Serviço de circulação e empréstimo 	Serviço que consiste no controlo de empréstimos, devoluções e gestão de reservas de obras.
	<ul style="list-style-type: none"> Serviço de referência: Apoio e orientação de 1ª linha 	Apoio aos utilizadores da Biblioteca na pesquisa e localização de obras.
	<ul style="list-style-type: none"> Uso dos espaços e equipamentos 	Consiste essencialmente na gestão da ocupação dos espaços de leitura, na gestão da utilização dos equipamentos de reprodução, na arrumação dos pisos da Biblioteca, em termos de mobiliário e de obras e no zelo pelo cumprimento das regras de conduta adequadas na Biblioteca.
Infoliteracia	<ul style="list-style-type: none"> Formação 	Consiste na oferta de ações de formação relacionadas com as áreas de recursos de informação.
	<ul style="list-style-type: none"> Serviço de referência: Apoio e orientação de 2ª linha 	Apoio especializado aos utilizadores da Biblioteca, em termos de pesquisas e acesso a recursos de informação, orientação para criação de referências bibliográficas.
	<ul style="list-style-type: none"> Serviço de apoio a utilizadores externos 	Apoio especializado a utilizadores externos à FEUP que se inscrevam na Biblioteca.
Tarefas de Gestão	<ul style="list-style-type: none"> Gestão dos pisos 	Consiste em tarefas de coordenação das equipas de recursos humanos que dão apoio aos pisos da Biblioteca.
	<ul style="list-style-type: none"> Gestão dos canais de comunicação 	Verificação permanente do endereço de correio eletrónico da Biblioteca, de Trouble Tickets, de Contatos telefónicos e Web Fóruns, respondendo às solicitações dos utilizadores de forma célere.
	<ul style="list-style-type: none"> Controlo de acessos 	Consiste em contabilizar as entradas e saídas da Biblioteca e registar quem a qualquer momento está na posse da Chave Mestra da Biblioteca ou do Cartão do Parque.
	<ul style="list-style-type: none"> Receção de Material 	Consiste em receber as encomendas de fornecedores que chegam ao piso 0 ou de obras compradas por elementos da FEUP, que as depositam na Biblioteca para tratamento técnico e ainda oferta de documentos.
	<ul style="list-style-type: none"> Controlo da Usabilidade e Acessibilidade de Documentos 	Registo de problemas na localização de documentos na Biblioteca e verificar e registar, caso seja necessário, o estado de conservação dos documentos devolvidos no piso 0.
	<ul style="list-style-type: none"> Registo de Problemas e Ocorrências 	Registo de problemas de funcionamento nos pisos da Biblioteca e problemas comportamentais dos utilizadores da Biblioteca.

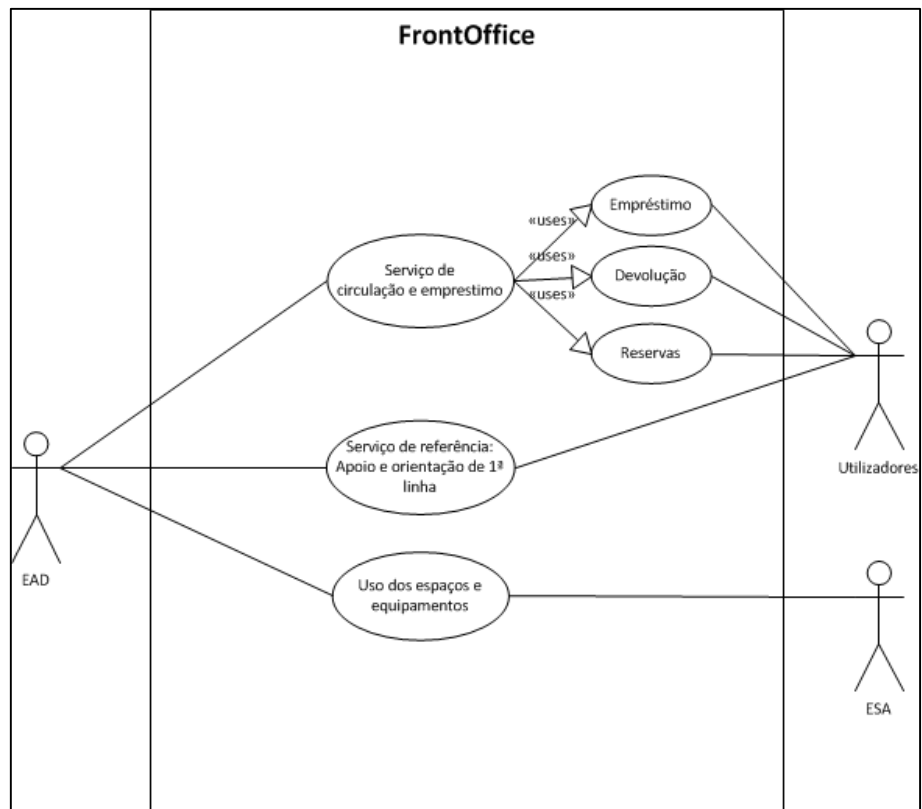


Figura C.1: Casos de Uso do processo FrontOffice

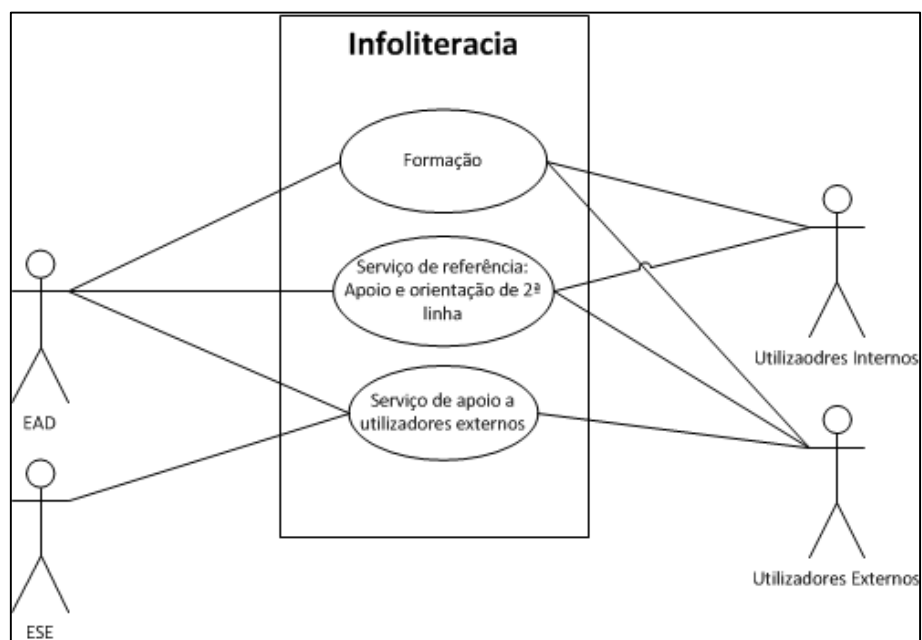


Figura C.2: Caso de Uso do processo Infoliteracia

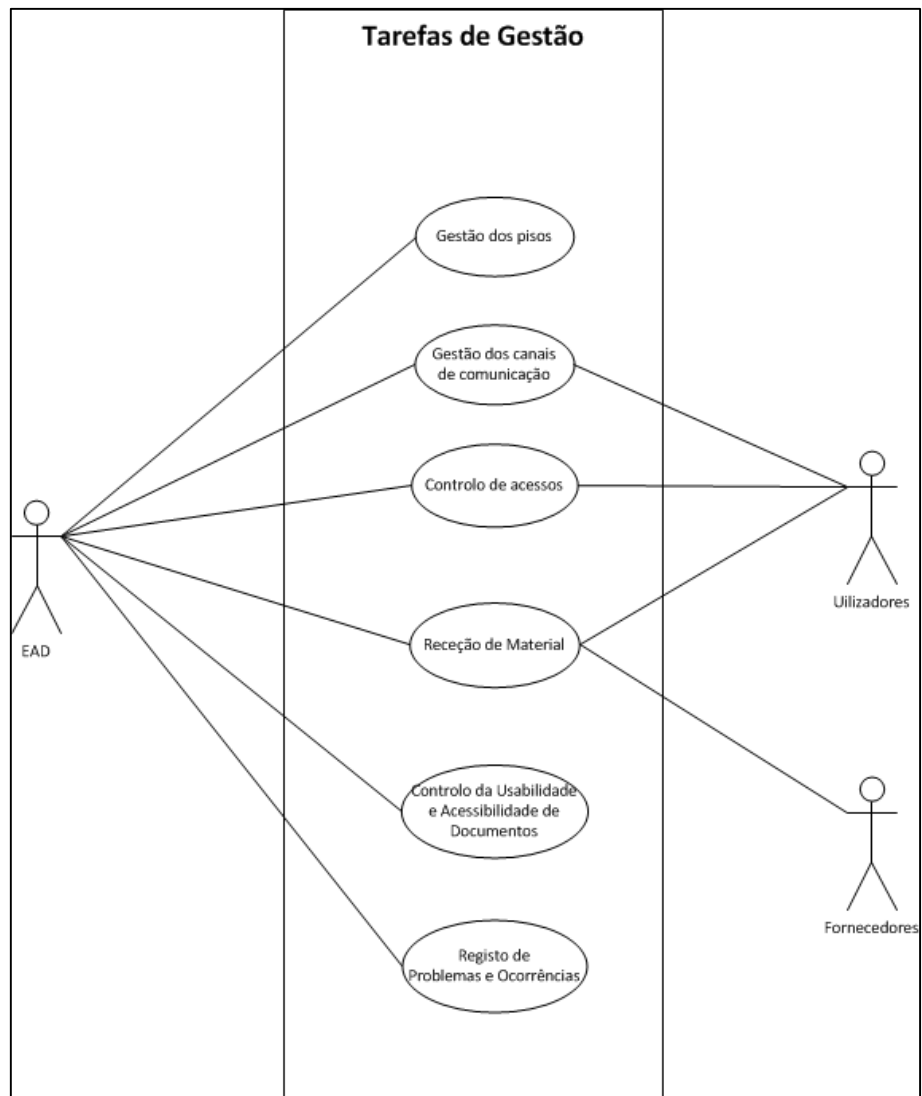


Figura C.3: Casos de Uso do processo Tarefas de Gestão

C.2 EQUIPA DE MEMÓRIA E ACESSO PERENE (EMAP)

Tabela 17: Sumário dos processos/atividades da equipa EMAP

Processos	Atividades	Descrição
Gestão Documental Física	<ul style="list-style-type: none"> Gestão de incorporações e alienações no Arquivo 	Acompanhamento, organização e triagem de recolhas e receção de ofertas, bem como de alienação de acervos.
	<ul style="list-style-type: none"> Gestão de incorporações e alienações na Biblioteca 	Acompanhamento da receção, organização e triagem de ofertas, bem como a alienação de obras.
	<ul style="list-style-type: none"> Avaliação do potencial museológico 	Determinar o valor museológico de material classificado para abate por serviços e departamentos. Normalmente isso acontece quando chega ao fim o ciclo de vida do material. Dependendo do seu valor, o material pode ser reaproveitado para o espólio museológico da FEUP.
	<ul style="list-style-type: none"> Gestão de espaços (depósitos e reservas) 	Gerir os espaços de depósito ou reserva, de acordo com a racionalização da instalação dos acervos.
Tratamento Técnico de Preservação	<ul style="list-style-type: none"> Intervenções técnicas de conservação 	Programação anual de intervenção técnica de conservação sobre os acervos patrimoniais da FEUP.
	<ul style="list-style-type: none"> Limpeza e higienização sistemática de documentos e espaços 	Programação anual de higienização e limpeza mecânica de documentos.
	<ul style="list-style-type: none"> Migração de informação entre suportes (Digitalização) 	Programa de transferência de informação entre suportes, nomeadamente através da digitalização de documentos em suporte convencional, papel.
	<ul style="list-style-type: none"> Monitorização ambiental 	Controlo dos fatores ambientais - Temperatura e Humidade Relativa - através de leituras e registos sistemáticos dos respetivos valores.
	<ul style="list-style-type: none"> Acondicionar e instalar os acervos 	Promover as práticas necessárias de acondicionamento e instalação para a boa preservação dos acervos.
Tarefas de Gestão	<ul style="list-style-type: none"> Regulamentação de áreas e processos de trabalho da gestão documental 	Elaboração de regulamentos e manuais de procedimentos.
	<ul style="list-style-type: none"> Acompanhamento e Avaliação das atividades realizadas 	Acompanhamento e gestão das atividades quotidianas da equipa.
	<ul style="list-style-type: none"> I&D&I 	Investigação e recolha de informação e experiências que promovam a qualidade e a inovação, e a adoção de boas práticas nos trabalhos desenvolvidos pela EMAP.

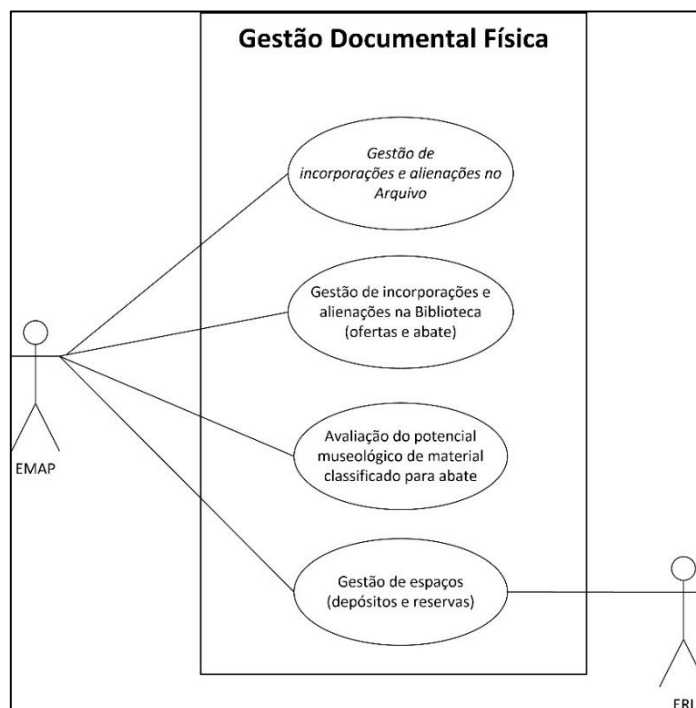


Figura C.4: Casos de Uso do processo Gestão Documental Física

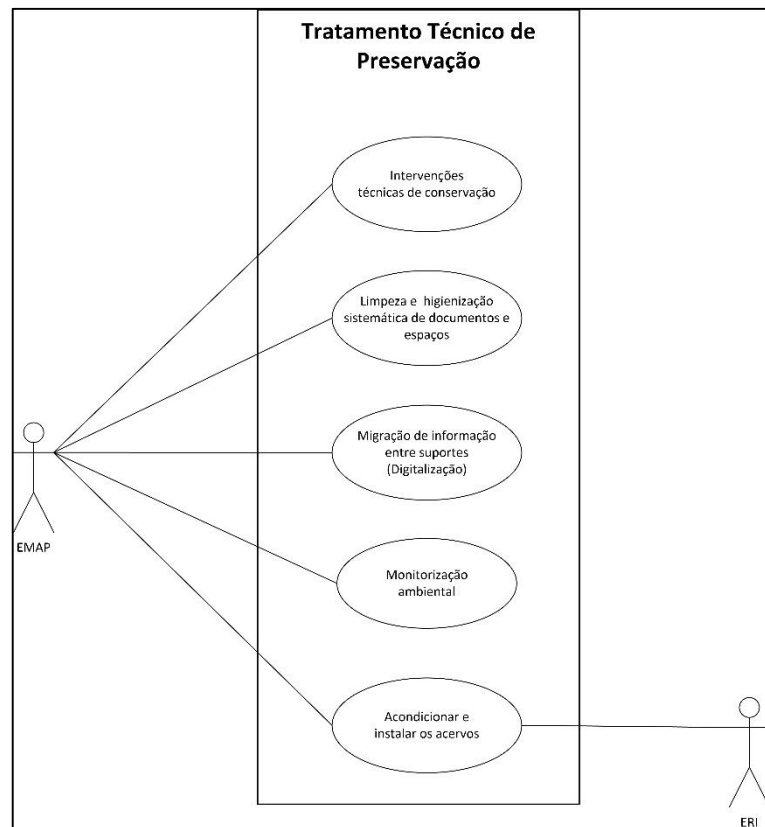


Figura C.5: Casos de Uso do processo Tratamento Técnico de Preservação

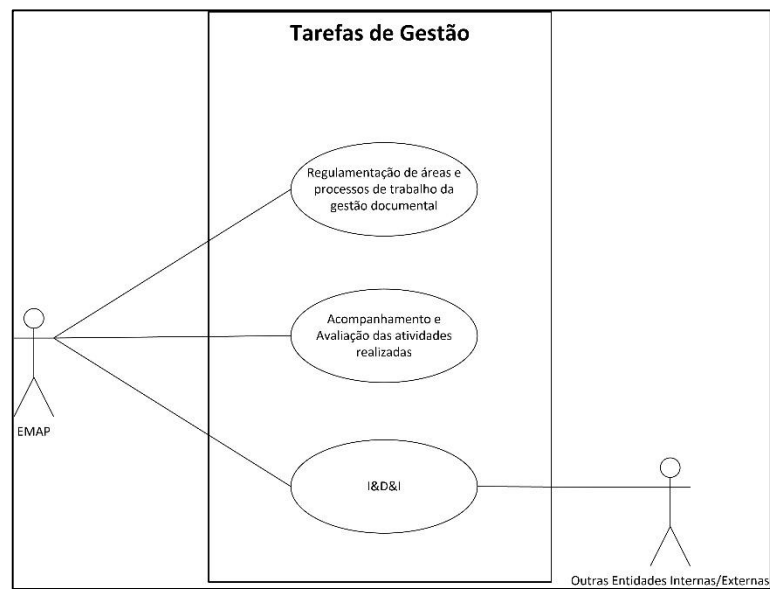


Figura C.6: Casos de Uso do processo Tarefas de Gestão (EMAP)

C.3 Equipa de Representação de Informação (ERI)

Tabela 18: Sumário dos processos/atividades da equipa ERI

Processos	Atividades	Descrição
Gestão documental	<ul style="list-style-type: none"> Tratamento documental 	Esta atividade consiste no Tratamento Preliminar (registo de posse e garantia de segurança), na descrição documental (inclui: catalogação, classificação, Indexação e cotação) e Finalização (inclui: colocação de cotas, colocação de proteções e etiquetas)
	<ul style="list-style-type: none"> Disponibilização de Documentos 	Avisos de disponibilidade, Envio para Novidades, Pisos, Armário de projetos, em linha, Requisição/Devolução de teses confidenciais, Preparação e disponibilização diária de jornais e revistas.
Tarefas de gestão	<ul style="list-style-type: none"> Planeamento e gestão 	Consiste nas seguintes tarefas: Controlo dos conteúdos documentais depositados na intranet do Serviço; Verificação permanentemente de endereços de correio, de Trouble Tickets e de Contatos telefónicos, respondendo às solicitações dos utilizadores de forma célere; Planeamento de relatórios de atividades e de avaliação; Participação em reuniões de coordenação, da ERI, específicas, com outras equipas/entidades.
	<ul style="list-style-type: none"> Controlo de qualidade 	Monitorização de procedimentos, revisão de procedimentos, registo de procedimentos em “Instruções de trabalho”. Registo de indicadores.
	<ul style="list-style-type: none"> Gerir a base de autoridade 	Controlo da entrada de termos nas bases de autoridade, bem como a verificação da qualidade dos registos de autoridade.
	<ul style="list-style-type: none"> Gerir as assinaturas correntes 	Esta atividade consiste no controlo das assinaturas das publicações periódicas subscritas pela FEUP. Inclui a verificação receção de material, a verificação do estado do material e a reclamação do atraso no fornecimento de fascículos esperados.

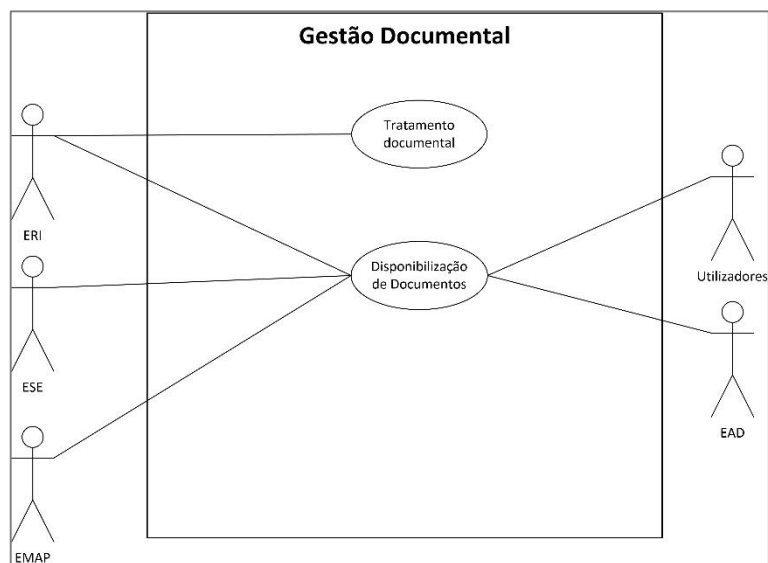


Figura C.7: Casos de Uso do processo Gestão Documental

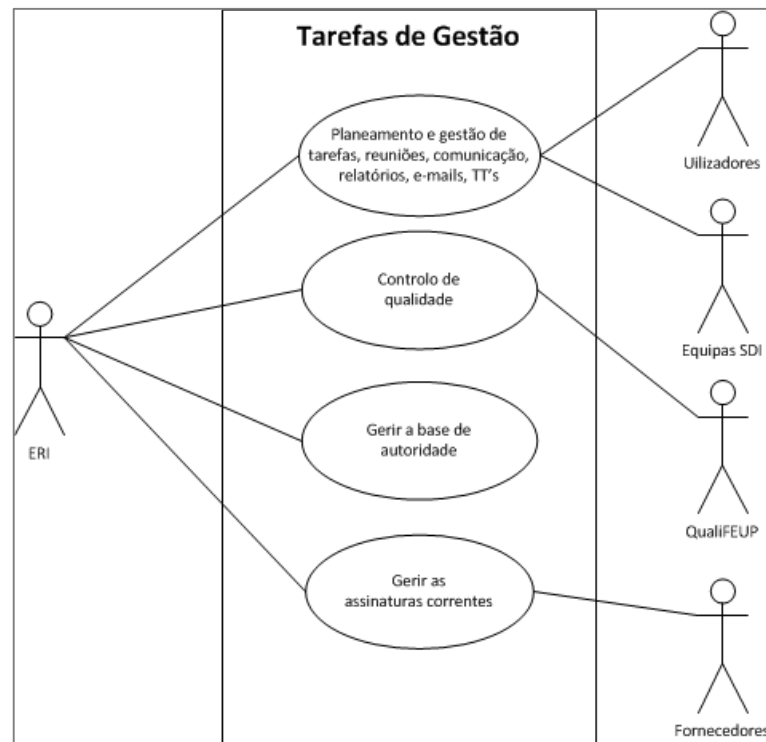


Figura C.8: Casos de Uso do processo Tarefas de Gestão

C.4 EQUIPA DE SUPORTE À ADMINISTRAÇÃO (ESA)

Tabela 19: Sumário dos processos/atividades da equipa ESA (EAD (2009); FEUP Edições (2009b, 2009a))

Processos	Atividades	Descrição
Coordenação	• Avaliação e diagnóstico	Supervisionar o progresso da execução dos objetivos estratégicos do SDI.
	• Planeamento	Planear as grandes linhas estratégicas orientadoras da atividade do SDI.
	• Decisão e orientação	Tomar decisões que garantam o cumprimento da estratégia e dos planos do Serviço.
	• Integração	Promover a integração das pessoas no Serviço e do Serviço na Comunidade que serve.
	• Comunicação e Feedback	Comunicação e Feedback
Secretariado Técnico	• Apoio administrativo, informativo e documental	Execução das tarefas inerentes à receção e distribuição de toda a documentação relativa ao SDI, minutar e dactilografar o expediente do SDI.
	• Assistir Diretora Serviços	Organização dos ficheiros e processos do SDI, assegurando a sua manutenção e atualização, bem como gerir e organizar o seu respetivo arquivo, na organização da pasta do expediente para o despacho da Diretora de Serviço e proceder de acordo com o seu despacho.
	• Despachar com Diretora Serviços	Preparação, elaboração e duplicação de documentação, convocação e marcação de reuniões e deslocações, sob a coordenação da Diretora de Serviços.
	• Gestão dados estatísticos SDI	Recolha, elaboração, disponibilização e distribuição dos dados estatísticos do SDI
	• Gestão Comunicacional	Atender, pessoalmente ou telefonicamente, as pessoas que procuram a Direção do SDI, orientá-las e prestar-lhes as informações necessárias, comunicar competentemente aos Superiores Hierárquicos os recados dos pretendentes, e, quando for o caso, encaminhá-los à Diretora de Serviços.
	• Gestão aplicação Regulamento Biblioteca	Elaboração de Relatórios Mensais e prestação de todo apoio administrativo – minutar, dactilografar e arquivar toda a documentação, trocada entre os Alunos, Direção SDI e Direção FEUP.
	• Colaborar na elaboração Plano de Atividades e Relatório de Atividades e Contas	Recolha junto dos diversos Serviços do SDI dos elementos a inserir no Plano de atividades e Relatório de Contas, bem como a produção gráfica.
Gestão de Infraestruturas	• Gestão de espaços	Garantir a preservação e bom funcionamento dos equipamentos, o mobiliário dos diferentes espaços dos edifícios afetos ao SDI, zelar pelo asseio e limpeza das Instalações do SDI (Biblioteca e Arquivo) e assegurar a aquisição dos bens e equipamentos necessários ao funcionamento do SDI.
	• Gestão de inventário	Garantir a atualização regular dos bens associados ao SDI.
	• Gestão Sistema Ambiental	Verificação regular do funcionamento do sistema de ambiental do SDI.
	• Gestão reserva de Salas	Validação dos pedidos de reserva de salas.
Gestão de Pessoas	• Gestão de competências	Garantir a clarificação das atribuições funcionais, definir os níveis de autonomia e de dependência.
	• Organização do trabalho	Garantir a clarificação das atribuições funcionais – com o respetivo <i>modus faciendi</i>
	• Gestão das obrigações e deveres	Garantir o conhecimento das obrigações legais e normativas internas; informar sobre modalidades de avaliação de cada pessoa e da equipa

Processos	Atividades	Descrição
	<ul style="list-style-type: none"> Gestão da avaliação 	Observar o nível do empenho e da qualidade do desempenho das pessoas, mantendo rotinas de avaliação das pessoas e das equipas
	<ul style="list-style-type: none"> Gestão de carreiras 	Implementar sistemas retributivos, compensatórios e meritórios, recorrendo aos dispositivos legais e a atividades de enriquecimento do currículo
	<ul style="list-style-type: none"> Gestão de iniciativas culturais e de encontro 	Garantir espaços de encontro, debate e de festa das pessoas que trabalham no SDI
Gestão Financeira	<ul style="list-style-type: none"> Elaboração de PADs 	Elabora os PAD's (Pedido de Autorização de Despesa) reunindo os documentos necessários para os mesmos e controla-os até à sua finalização.
	<ul style="list-style-type: none"> Elaboração de mapas de despesas e receitas 	Elaboração de documentos e mapas de despesas e receitas para apoio à tomada de decisões da Direção do SDI
	<ul style="list-style-type: none"> Elabora relatórios de execução orçamental 	Elabora relatórios de execução orçamental para o registo e controlo rigoroso de todas as despesas e receitas associadas ao SDI e a todos os centros de custos associados ao mesmo
	<ul style="list-style-type: none"> Controla o fundo de maneo associado ao SDI 	Controla o fundo de maneo associado ao SDI
Gestão de Aquisições	<ul style="list-style-type: none"> Verificação das necessidades de aquisição/contratação 	Verificação periódica dos produtos necessários a adquirir/renovar, por exemplo: livros, material de economato, renovações de serviços, etc.
	<ul style="list-style-type: none"> Pesquisa de mercado 	Pesquisa de mercado para obter o melhor produto a melhores preços.
	<ul style="list-style-type: none"> Gestão de encomendas 	Contactos com fornecedores para orçamentos e aquisições.
	<ul style="list-style-type: none"> Elaboração de cadernos de encargos 	Elaboração de cadernos de encargos e reunião dos documentos de apoio às aquisições.
	<ul style="list-style-type: none"> Cabimentação 	Alocação de verbas a aquisições.
	<ul style="list-style-type: none"> Autorização de despesa 	Verificação da disponibilidade de orçamento e autorização superior para posterior aquisição
Gestão da FEUP Edições	<ul style="list-style-type: none"> Avaliação e aprovação de propostas de publicação 	Assegurar que as obras publicadas pela FEUP Edições respeitam os padrões de qualidade técnico-científica estabelecidos pela FEUP.
	<ul style="list-style-type: none"> Preparação 	Tomar todas as medidas necessárias para o início da produção da obra. Comporta as seguintes atividades: Planear trabalhos, Contratar publicação, Fornecer dados, informação e elementos de carácter editorial e/ou gráfico.
	<ul style="list-style-type: none"> Produção 	Esta atividade corresponde à fase de produção da obra e que culmina com a receção da obra (o produto final).
	<ul style="list-style-type: none"> Comercialização 	Assegurar que as obras publicadas pela FEUP edições estejam disponíveis no mercado editorial, de acordo com a política de comercialização definida pela FEUP edições.
	<ul style="list-style-type: none"> Comunicação e divulgação 	Assegurar que a obra seja do conhecimento dos potenciais clientes. Compreende as seguintes tarefas: Divulgação interna, Divulgação externa, Divulgação orientada e Divulgação no sitio WEB
Gestão da Qualidade	<ul style="list-style-type: none"> Identificação das expectativas dos utilizadores 	Desenvolver métodos e instrumentos de escuta intra e extra- organizacional no sentido de conhecer e compreender as necessidades e expectativas dos utilizadores.
	<ul style="list-style-type: none"> Análise do desempenho do Serviço 	Tratar e analisar os dados recolhidos referentes ao desempenho do serviço.
	<ul style="list-style-type: none"> Análise dos dados recolhidos 	Tratar e analisar os dados recolhidos referentes à perceção que o utilizador tem do serviço. Tratar e analisar dados recolhidos referentes ao ambiente intraorganizacional.
	<ul style="list-style-type: none"> Propostas de medidas corretoras 	Propor medidas corretoras de deficiências encontradas.

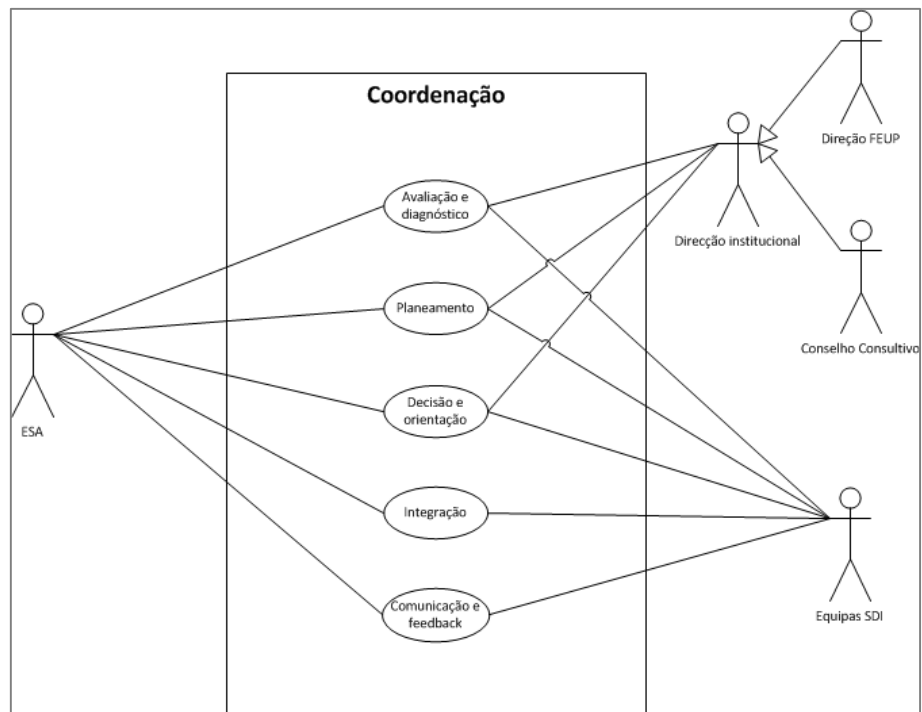


Figura C.9: Casos de Uso do processo Coordenação

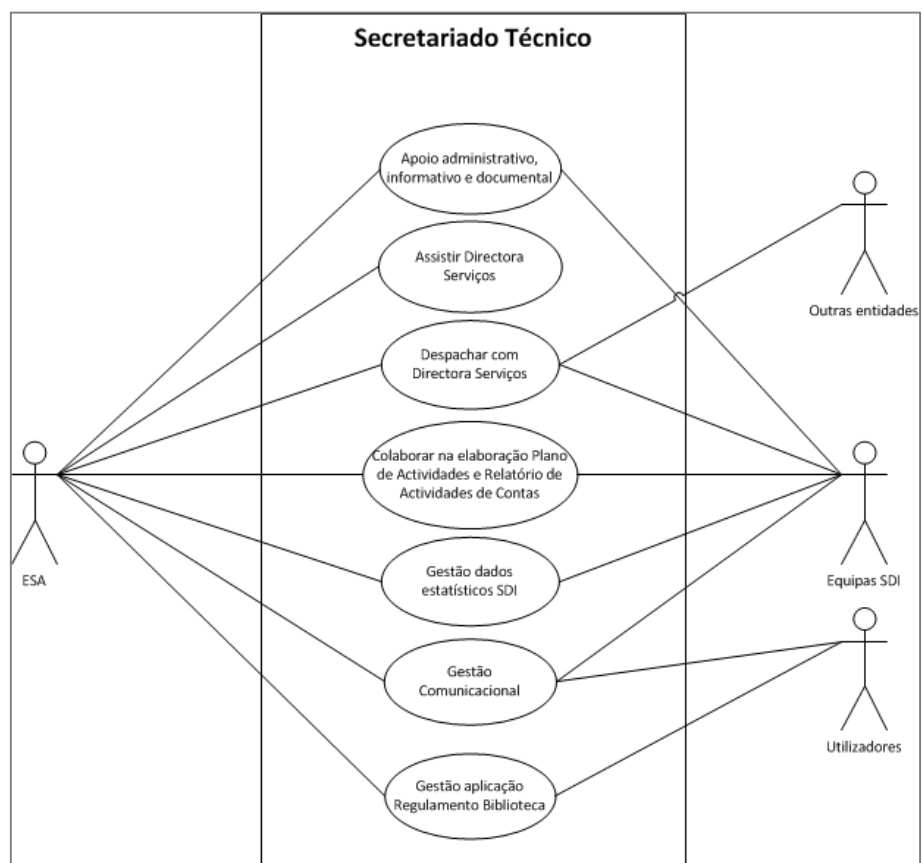


Figura C.10: Casos de Uso do processo Secretariado Técnico

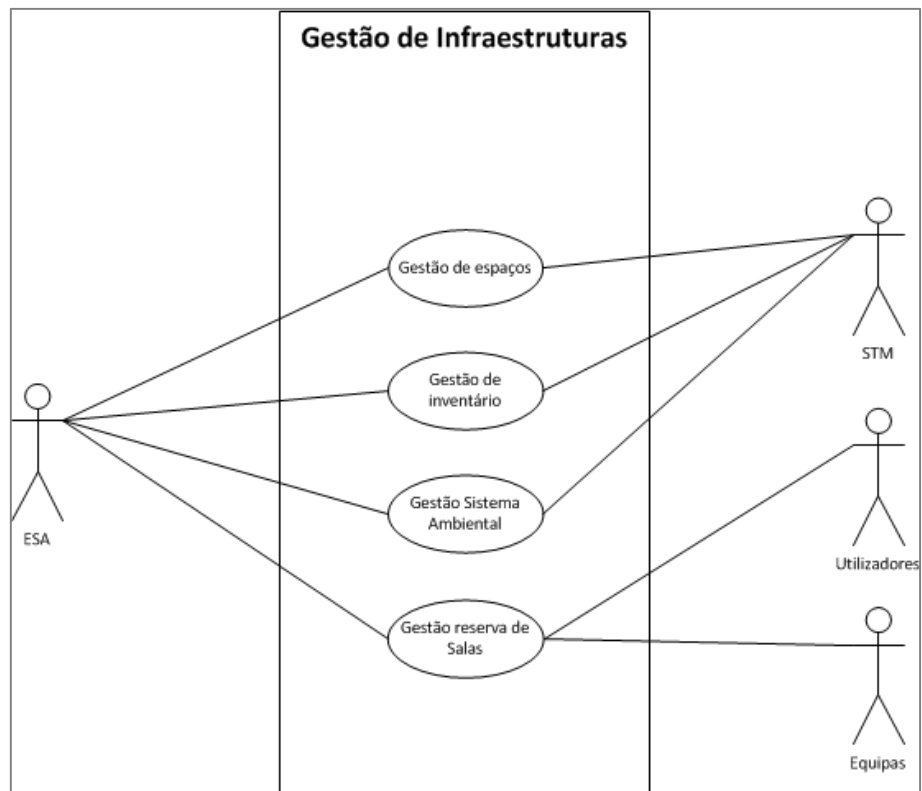


Figura C.11: Casos de Uso do processo Gestão de Infraestruturas

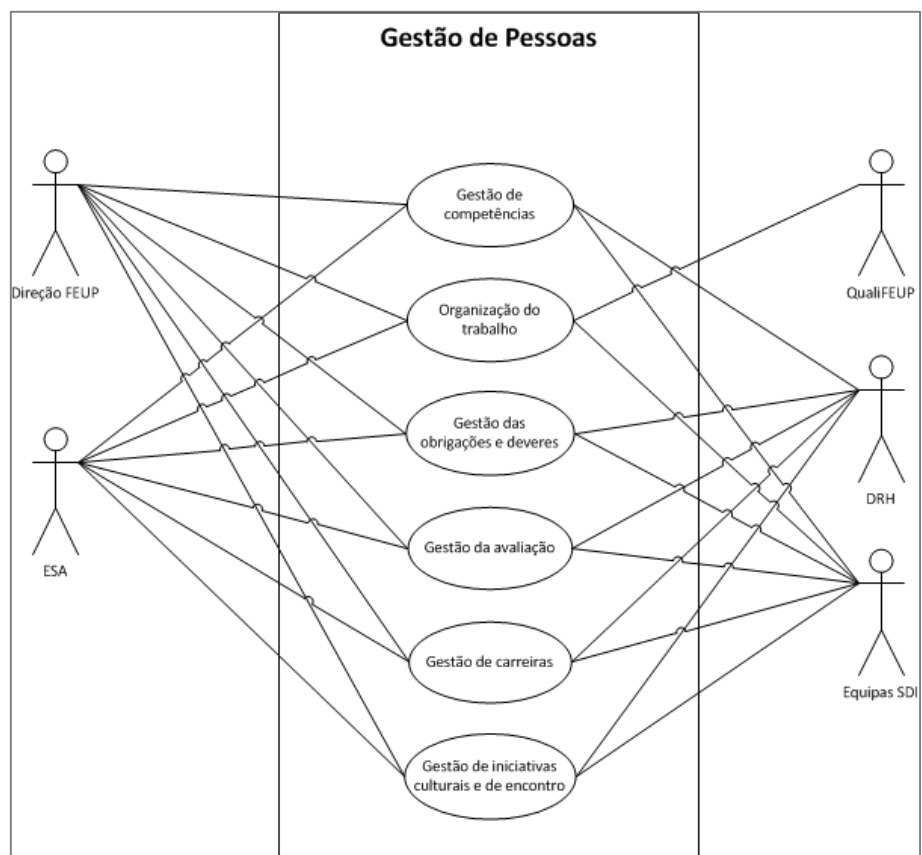


Figura C.12: Casos de Uso do processo Gestão de Pessoas

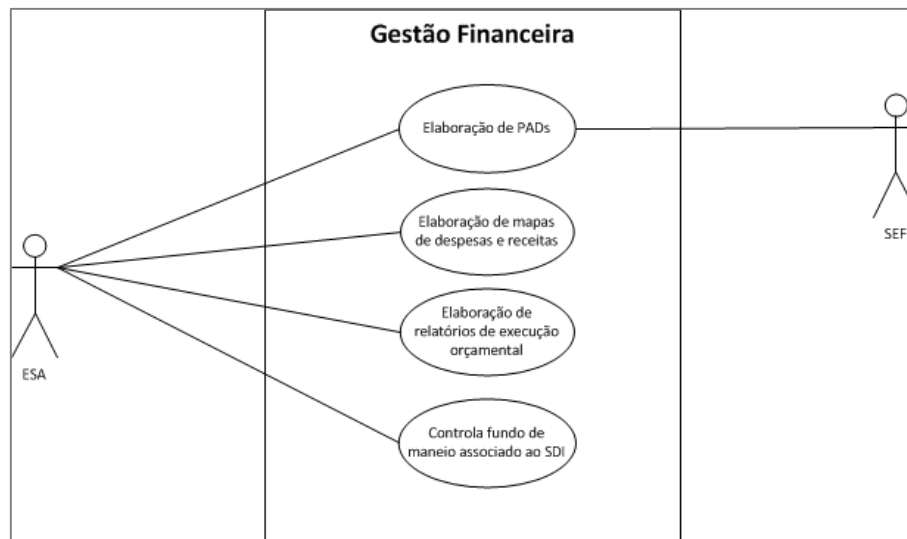


Figura C.13: Casos de Uso do processo Gestão Financeira

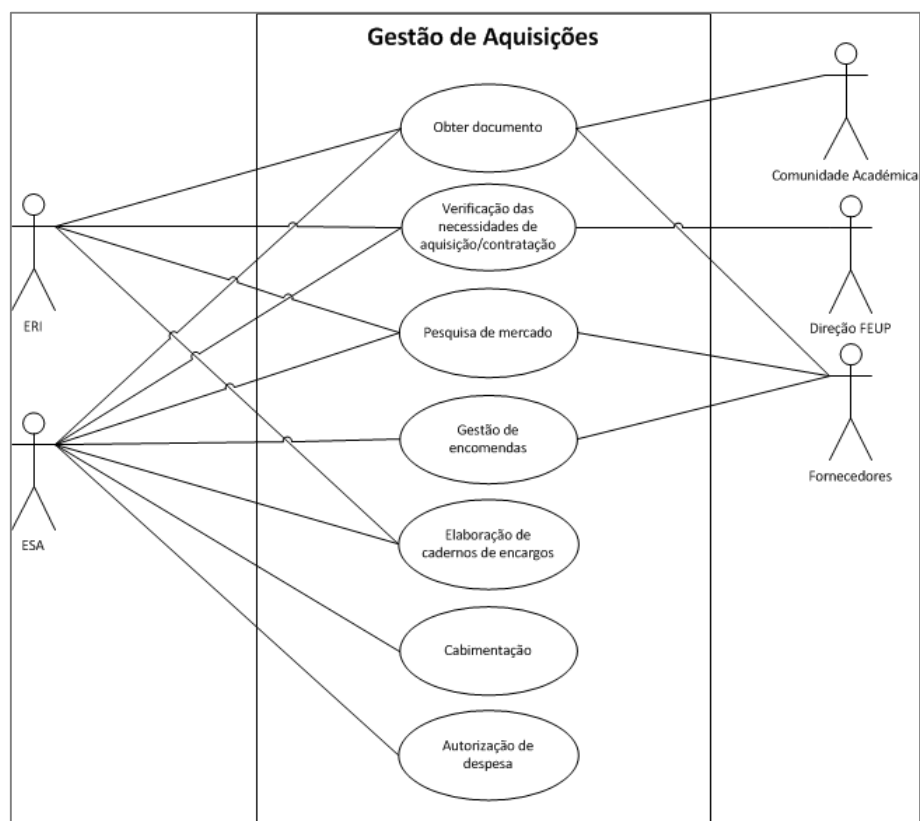


Figura C.14: Casos de Uso do processo Gestão de Aquisições

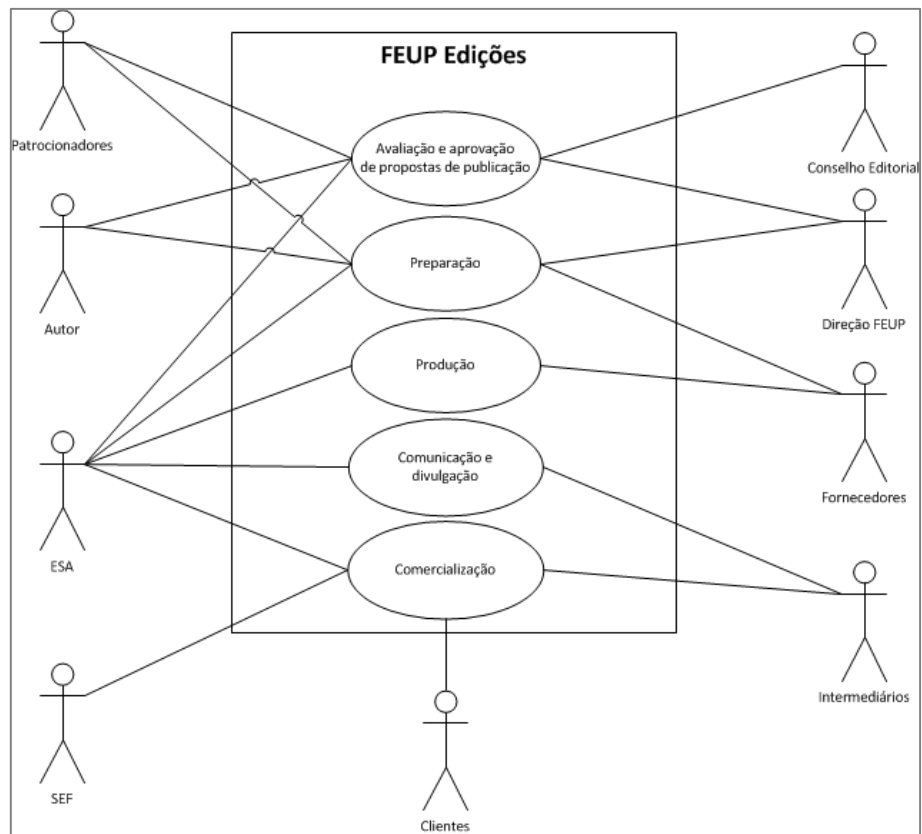


Figura C.15: Casos de Uso do processo FEUP Edições

C.5 Equipa de Serviços Eletrónicos (ESE)

Tabela 20: Sumário dos processos/atividades da equipa ESE

Processos	Atividades	Descrição
Produção de Documentos	<ul style="list-style-type: none"> Produção de versões eletrónicas de documentos 	Consiste em a partir de documentos em formatos tradicionais, essencialmente papel, produzir cópias eletrónicas no formato mais adequado à sua disponibilização e preservação (ex: jpeg, tiff, pdf, vídeos, flv).
	<ul style="list-style-type: none"> Produção de livros eletrónicos³ 	Produção para a FEUP Edições de livros eletrónicos a partir dos conteúdos elaborados pelos autores.
Projetos	<ul style="list-style-type: none"> Desenvolvimento de novas funcionalidades 	Avaliação e implementação de plataformas e sistemas de suporte às atividades do SDI e no desenvolvimento de novos serviços.
Gestão de Serviços WEB	<ul style="list-style-type: none"> Criação e Gestão de Serviços Web 	Desenho, implementação e gestão de aplicações WEB de apoio às atividades do SDI.
	<ul style="list-style-type: none"> Web Design 	Desenho de interfaces Gráficas WEB para as aplicações anteriormente mencionadas e na produção de materiais de suporte à comunicação do SDI.
Gestão da infraestrutura tecnológica	<ul style="list-style-type: none"> Gestão e desenvolvimento das aplicações de suporte 	Gestão corrente das aplicações de suporte às atividades do SDI.
	<ul style="list-style-type: none"> Manutenção dos servidores e computadores em uso no SDI 	Gerir os servidores e computadores do SDI de forma a garantir a sua adequação aos requisitos necessários à prestação dos serviços do SDI aos utilizadores.

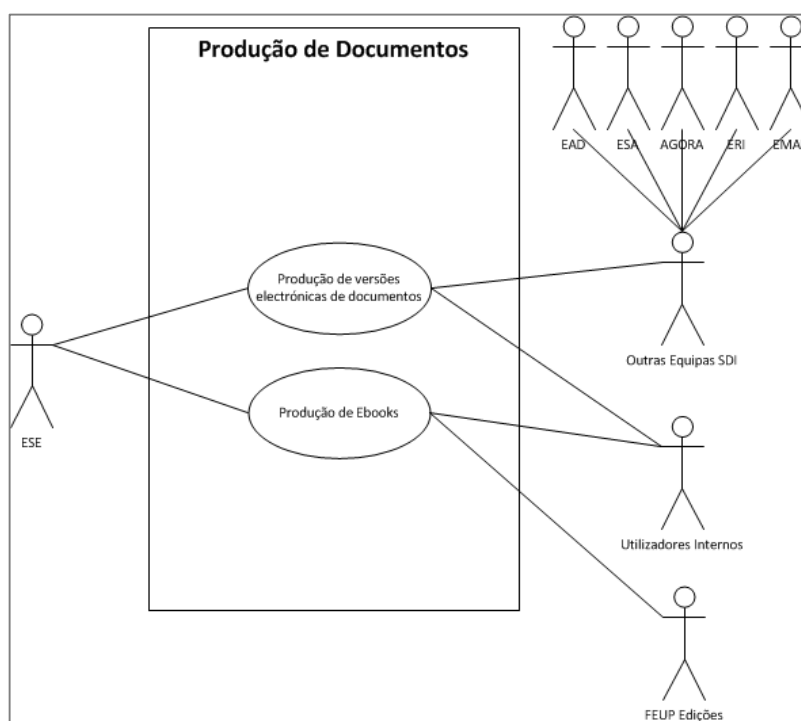


Figura C.16: Casos de Uso do processo Produção de Documentos

³ Do Inglês: EBook

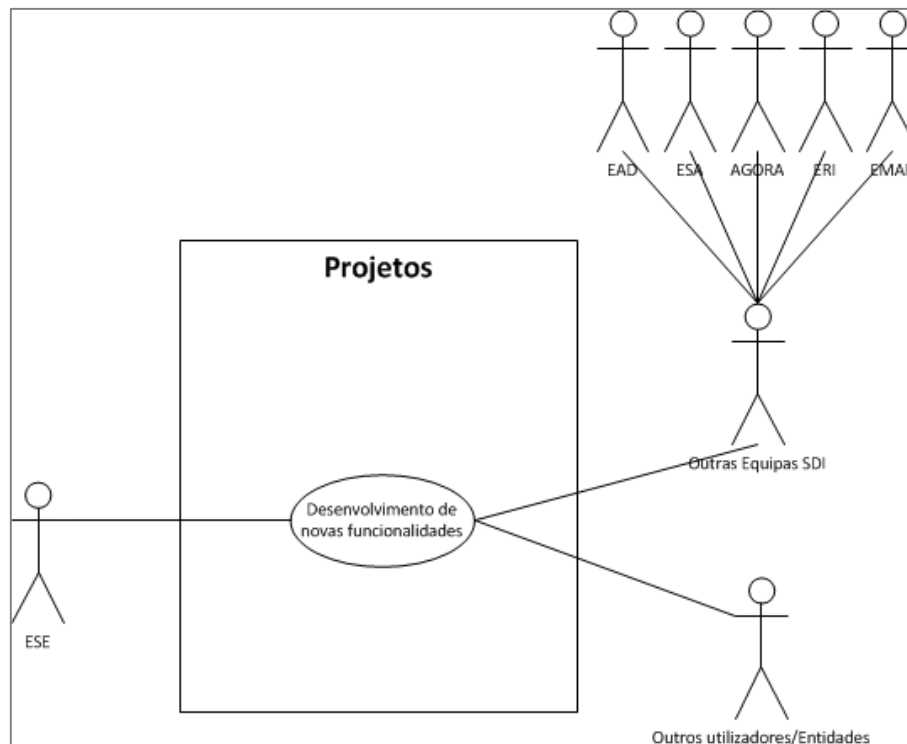


Figura C.17: Casos de Uso do processo Projetos

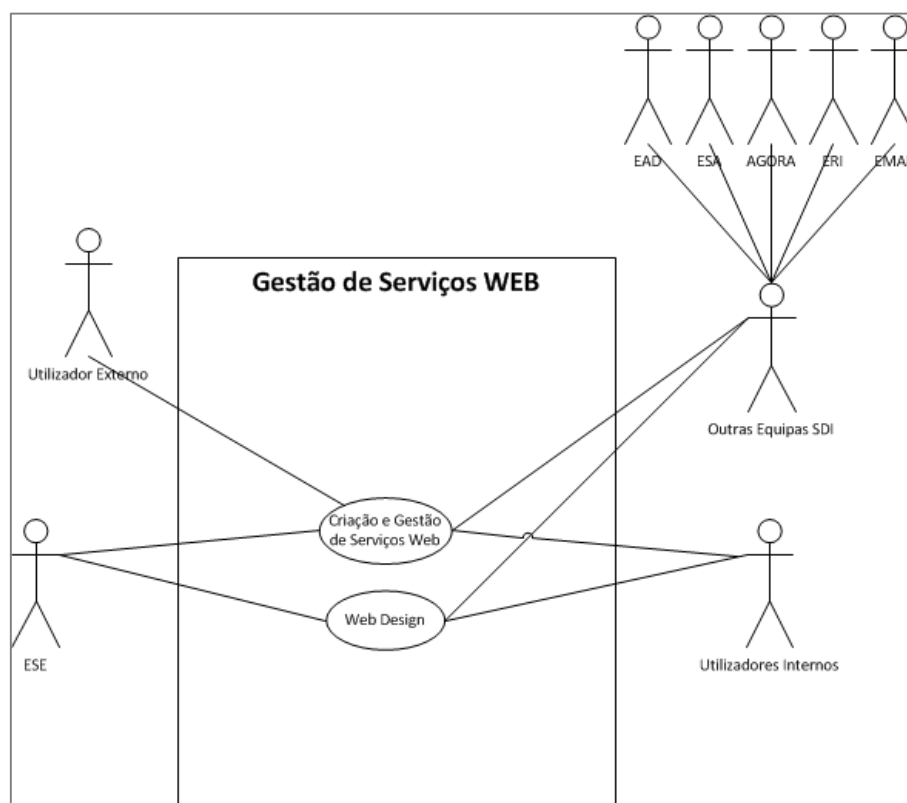


Figura C.18: Casos de Uso do processo Gestão de Serviços Web

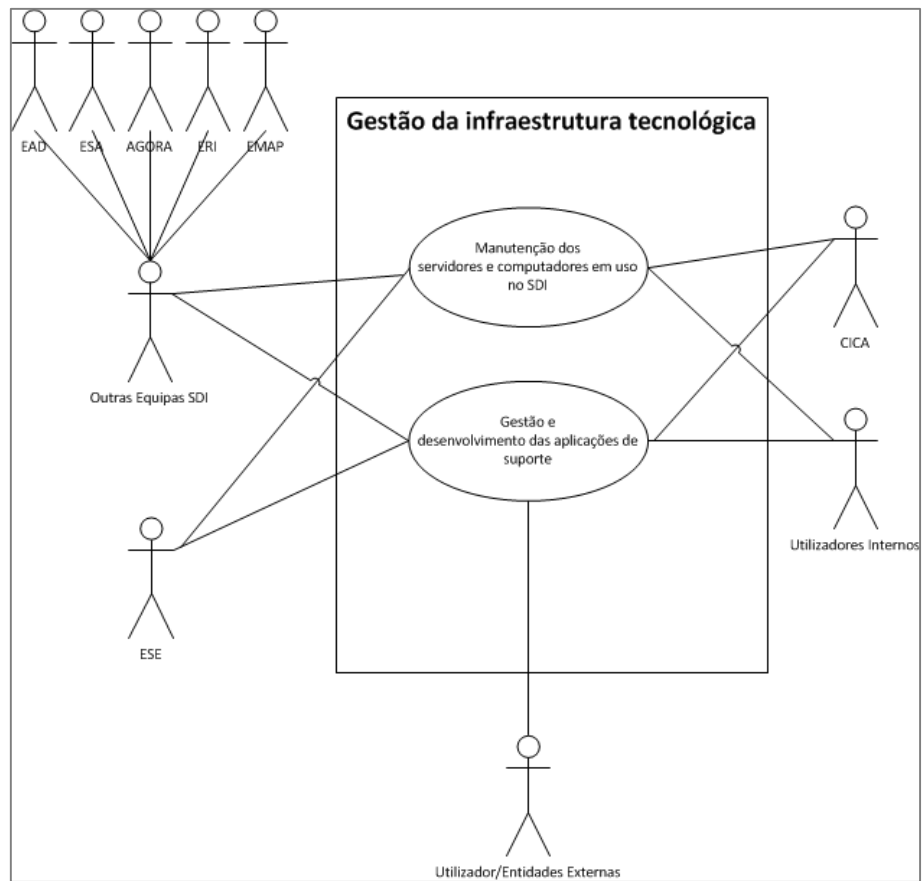


Figura C.19: Casos de Uso do processo Gestão da infraestrutura tecnológica

Anexo D. Matriz de processos e fontes de dados

D.1 Matriz de processos e fontes de dados da equipa EAD

Processos	Atividades	Fontes	Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Fontes ad hoc
Front office	Serviço de circulação e empréstimo		●				●	●															●
	Serviço de referência: Apoio e orient		●	●	●	●			●	●		●											●
	Uso dos espaços e equipamentos		●								●												
Infoliteracia	Formação		●	●	●	●			●	●		●	●	●	●								●
	Serviço de referência: Apoio e orient		●	●	●	●			●	●	●	●	●	●	●						●		●
	Serviço de apoio a utilizadores exter		●																	●			
Tarefas de gestão	Gestão dos pisos																						●
	Gestão dos canais de comunicação										●										●		
	Controlo de acessos																					●	●
	Receção de Material																					●	●
	Controlo da Usabilidade e Acessibili	●									●	●										●	●
	Registo de Problemas e Ocorrências										●	●											

D.2 Matriz de processos e fontes de dados da equipa EMAP

Processos	Atividades	Fontes	Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Fontes ad hoc
Gestão Documental Física	Gestão de incorporações e alienações no Arq							●													●		●
	Gestão de incorporações e alienações na Biblio		●																		●		●
	Avaliação do potencial museológico de materia										●										●		●
	Gestão de espaços (depósitos e reservas)		●					●				●									●		●
Tratamento Técnico de Preservação	Intervenções técnicas de conservação							●				●									●		●
	Limpeza e higienização sistemática de docume											●									●		●
	Migração de informação entre suportes (Digit		●	●			●	●				●	●	●							●		●
	Monitorização ambiental											●									●		●
Tarefas de Gestão	Acondicionar e instalar os acervos											●									●		●
	Regulamentação de áreas e processos de trab											●									●		●
	Acompanhamento e Avaliação das atividades I&D&I											●									●		●

D.3 Matriz de processos e fontes de dados da equipa ESE

Processos	Atividades	Fontes	Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Fontes ad hoc
Produção de Documentos	Produção de versões electrónicas		●	●							●	●											●
	Produção de Ebooks																						●
Projetos	Desenvolvimento de novas funcio										●								●				●
Gestão de Serviços WEB	Criação e Gestão de Serviços Web								●				●	●	●	●							●
	Web Design												●	●									●
Gestão da infraestrutura	Gestão e desenvolvimento das apl		●	●	●	●	●	●				●	●	●	●	●			●	●			
	Manutenção dos servidores e com		●	●	●	●	●	●			●		●	●	●	●			●	●	●		

D.4 Matriz de processos e fontes de dados da equipa AGORA

Processos	Atividades	Fontes	Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Dropbox	Fontes ad hoc
Coordenação	Conceber projeto											●												●
	Gerir projeto																							●
Interpretação	I&D&I											●									●			●
	Recolher Informação		●									●									●			●
	Conceber proposta de interpretação						●					●											●	●
Eventos	Programar											●									●		●	●
	Produzir											●									●		●	

D.5 Matriz de processos e fontes de dados da equipa ESA

Processos	Atividades	Fontes	Aleph	Digitool	Metalib	SFX	InArte	GISA	Recursos eletrónicos	Moodle	SIGARRA	Sharepoint	Seagull	Drupal	MediaWiki (GAP)	Open Journal System	GRH	QualiFEUP	Site FEUP Edições	Redmine	Email	Google Books	Fontes ad hoc
Coordenação	Avaliação e diagnóstico										●	●					●				●		
	Planeamento										●	●					●				●		
	Decisão e orientação										●	●					●				●		
	Integração										●	●					●				●		
	Comunicação e feedback										●	●	●	●							●		
Secretariado Técnico	Apoio administrativo, informativo										●	●					●				●		●
	Assistir Directora Serviços										●	●									●		●
	Despachar com Directora Serviços										●	●									●		●
	Gestão dados estatísticos SDI	●	●	●	●	●	●	●				●									●		●
	Gestão Comunicacional										●	●					●				●		●
	Gestão aplicação Regulamento Bib											●									●		●
	Colaborar na elaboração Plano de											●									●		●
Gestão de Infra-estruturas	Gestão de espaços										●	●									●		●
	Gestão de inventário										●	●									●		●
	Gestão Sistema Ambiental										●	●									●		●
	Gestão reserva de Salas										●	●									●		●
Gestão de Pessoas	Gestão de competências										●	●					●						
	Organização do trabalho										●	●						●					
	Gestão das obrigações e deveres										●						●		●				
	Gestão da avaliação										●						●						
	Gestão de carreiras										●						●				●		
	Gestão de iniciativas culturais e de											●	●	●							●		●
Gestão Financeira	Elaboração de PADs										●												●
	Elaboração de mapas de despesas										●												●
	Elabora relatórios de execução org																					●	●
	Controla o fundo de manei										●												●
Gestão de Aquisições	Obter documento	●									●										●		●
	Verificação das necessidades de ad	●									●												●
	Pesquisa de mercado	●																					●
	Gestão de encomendas	●									●												●
	Elaboração de cadernos de encarg																						●
	Cabimentação	●									●												●
	Autorização de despesa										●												●
Gestão da FEUPedições	Avaliação e aprovação de proposta											●									●		●
	Preparação											●									●		●
	Produção											●									●		●
	Comercialização										●	●						●			●		●
	Comunicação e divulgação										●	●	●	●				●			●	●	●

Anexo E. Matriz de barramento dimensional

Data mart	Estrela	Dimensões																																				
		Documento	Recurso de pesquisa	Site temático	Aplicação de suporte	Funcionalidade	Reclamação	Termo de autoridade	Index_grp ⁴	Colaborador	Utilizador	Bibliografia UC	Unidade Curricular	Curso	UO	Categoria Pedido Apoio	Sala	Estante depósito	Equipamento	Incidente	Projeto cooperação	Projeto	Evento	Tarefa	Instrução de trabalho	Plano de atividade	Orçamento	Rubrica	CCO	Produto	PAD	Fornecedor	Humidade (HR)	Temperatura	Luminosidade (lx)	Tempo		
Gestão Documental	Documentos novos	●							●																											●		
	Documentos alterados	●							●																											●		
	Exemplares novos	●							●																											●		
	Exemplares alterados	●							●																											●		
	Receção documentos para tratamento	●							●																											●		
	Registos autoridade novos						●		●																											●		
	Registos autoridade alterados						●		●																											●		
	Indexação	●						●																													●	
Tarefas de Gestão ERI	Tratamento técnico	●							●																												●	
	Controlo de qualidade	●							●															●													●	
	Outras colaborações								●													●															●	
Frontoffice	Inscrições									●																											●	
	Circulação	●							●	●				●	●																						●	
	Multas	●							●	●																											●	
	Obras extraviadas	●							●																												●	
	Requisições Arquivo															●																					●	
	Consultas arquivo acompanhadas								●	●					●																						●	
	Consultas arquivo desacompanhadas														●																						●	
	Pedido de apoio 1.ª linha								●	●						●																					●	
	Ocupação Pisos																																					●
	Uso Fotocopiadoras																																					●
	Reserva Gabinetes								●	●																											●	

⁴ A dimensão “Index_grp” na realidade não é uma dimensão. Trata-se de uma tabela de ponte que faz a correspondência entre documentos e o conjunto dos assuntos que lhe foram atribuídos pela indexação.

Data mart	Estrela	Dimensões																																		
		Documento	Recurso de pesquisa	Site temático	Aplicação de suporte	Funcionalidade	Reclamação	Termo de autoridade	Index_grp	Colaborador	Utilizador	Bibliografia UC	Unidade Curricular	Curso	UO	Categoria Pedido Apoio	Sala	Estante depósito	Equipamento	Incidente	Projeto cooperação	Projeto	Evento	Tarefa	Instrução de trabalho	Plano de atividade	Orçamento	Rubrica	CCO	Produto	PAD	Fornecedor	Humidade (HR)	Temperatura	Luminosidade (lx)	Tempo
Infoliteracia	Formação								●				●	●								●														●
	Participação								●	●				●																						●
	Pedido de apoio 2.ª linha									●	●				●																					●
	Utilizadores externos								●	●																										●
	Sessões		●																																	●
	Pesquisas		●																																	●
Tarefas de Gestão EAD	Visualizações	●	●																																	●
	TT EAD								●	●					●																					●
	Entradas Bib.																																			●
	Recepção Material	●							●	●					●																					●
	Extravios	●							●																											●
	Danos	●							●																											●
Gestão Documental Física	Problemas comportamentais									●																										●
	Sanções									●																										●
	Oferta Documentos	●								●																										●
	Ofertas Museu	●								●																										●
	Incorporações no Arquivo	●							●	●					●																					●
	Incorporações na Biblioteca	●							●	●																										●
	Ocupação depósitos	●													●		●	●																	●	
	Transferências no Arquivo	●							●						●																					●
	Alienações no Arquivo	●							●						●		●	●																	●	
Alienações na Biblioteca	●							●						●		●	●																	●		
Gestão Documental Física	Instalação Documentos	●							●																											●
	Instalação Museu	●							●																											●

Data mart	Estrela	Dimensões																																							
		Documento	Recurso de pesquisa	Site temático	Aplicação de suporte	Funcionalidade	Reclamação	Termo de autoridade	Index_grp	Colaborador	Utilizador	Bibliografia UC	Unidade Curricular	Curso	UO	Categoria Pedido Apoio	Sala	Estante depósito	Equipamento	Incidente	Projeto cooperação	Projeto	Evento	Tarefa	Instrução de trabalho	Plano de atividade	Orçamento	Rubrica	CCO	Produto	PAD	Fornecedor	Humidade (HR)	Temperatura	Luminosidade (lx)	Tempo					
Tratamento Técnico de Preservação	Monitorização Ambiental																●																●	●	●	●					
	Higienização Documentos	●							●																											●	●				
	Higienização Museu	●							●																												●	●			
	Digitalização	●							●																												●	●			
	Encadernação	●								●																												●	●		
Produção de documentos	Produção	●							●																													●	●		
	Depósito	●							●																													●	●		
	Disponibilização	●																																				●	●		
	Ebooks	●							●																													●	●		
	Satisfação Ebooks	●									●													●														●	●		
Projetos	Desenvolvimento Projetos				●					●												●		●														●	●		
Gestão de Serviços WEB	Satisfação Sites			●						●													●		●													●	●		
	Desenvolvimento Webdesign								●													●		●														●	●		
	Comportamento			●																																			●	●	
Infraestrutura tecnológica	Disponibilidade				●																																		●	●	
	Reclamações ESE						●				●																												●	●	
	Manutenção				●					●																													●	●	
	Desenvolvimento aplicações				●	●				●																													●	●	
	Satisfação desenvolvimentos					●					●																												●	●	
Coordenação AGORA	Conceção								●													●	●																●	●	
	Desenvolvimento AGORA								●													●	●	●															●	●	
Interpretação	Interpretação	●							●														●	●																●	●

Data mart	Estrela	Dimensões																																					
		Documento	Recurso de pesquisa	Site temático	Aplicação de suporte	Funcionalidade	Reclamação	Termo de autoridade	Index_grp	Colaborador	Utilizador	Bibliografia UC	Unidade Curricular	Curso	UO	Categoria Pedido Apoio	Sala	Estante depósito	Equipamento	Incidente	Projeto cooperação	Projeto	Evento	Tarefa	Instrução de trabalho	Plano de atividade	Orçamento	Rubrica	CCO	Produto	PAD	Fornecedor	Humidade (HR)	Temperatura	Luminosidade (lx)	Tempo			
Eventos	Participação																						●														●		
	Satisfação AGORA									●													●														●		
	Relações														●								●														●		
	Impacto	●																						●														●	
Coordenação ESA	Execução PA																							●		●												●	
	Visitas									●	●																											●	
	Instruções trabalho novas									●															●													●	
	Instruções trabalho corrigidas									●															●													●	
	Projetos cooperação									●											●																	●	
Reclamações ESA						●				●																												●	
Secretariado Técnico	Divulgação																						●															●	
Gestão de Infraestruturas	Condições ambientais									●							●																	●	●				●
	Gestão espaços									●							●			●																		●	
	Gestão equipamentos									●							●		●	●																		●	
	Inventário									●							●		●																			●	
Gestão de Pessoas	Assiduidade									●																												●	
	Reuniões									●													●															●	
Gestão Financeira	Movimentos										●																●	●	●	●	●	●							●
Gestão de Aquisições	Bibliografia UC											●	●	●	●																							●	
	Aquisições	●								●	●																	●		●		●	●						●
	EIB	●								●	●																		●		●	●							●
	Comparativo																													●		●						●	
	Reposições	●								●	●																												●
Gestão da FEUP Edições	Vendas	●	●																																				●
	Preparação	●																						●														●	
	Produção	●																						●														●	

Anexo F. Dicionário de factos

F.1 Factos dos *data marts* da equipa ERI

Estrela	Documentos novos	Versão	1.1	Data	07/02/2013
Granularidade	Registo de um documento criado				
Dimensões					
Documento	Documento				
Colaborador	Colaborador				
Tempo	Tempo				
Medidas					
Quantidade	N.º de registos novos criados na Biblioteca, Museu ou Arquivo				

Estrela	Documentos alterados	Versão	1.2	Data	21/03/2013
Granularidade	Registo de um documento alterado				
Dimensões					
Documento	Documento				
Colaborador	Colaborador				
Tempo	Tempo				
Medidas					
Quantidade	N.º de registos alterados na Biblioteca, Museu ou Arquivo				

Estrela	Exemplares novos	Versão	1.2	Data	21/03/2013
Granularidade	Exemplar de um documento criado				
Dimensões					
Documento	Documento				
Colaborador	Colaborador				
Tempo	Tempo				
Medidas					
Quantidade	Número de exemplares criados na Biblioteca, Museu ou Arquivo				

Estrela	Exemplares alterados	Versão	1.2	Data	21/03/2013
Granularidade	Exemplar de um documento alterado				
Dimensões					
Documento	Documento				
Colaborador	Colaborador				
Tempo	Tempo				
Medidas					
Quantidade	Número de exemplares alterados na Biblioteca, Museu ou Arquivo				

Estrela	Receção material	Versão	1.2	Data	07/03/2013
Granularidade	Receção de exemplares de um documento para tratamento técnico				
Dimensões					
Colaborador	Quem recebe o documento				
Tipo Entrega (DD)	Encomenda, Entrega, Doação				
Tipo Material (DD)	Livro, Tese, Periódico, CD-DVD,...				
Tempo	Data da receção do documento				
Medidas					
Quantidade	Número de exemplares rececionados, por documento				

Estrela	Autoridades novas	Versão	1.2	Data	05/04/2013
Granularidade	Registo de autoridade novo inserido no sistema da Biblioteca, Museu ou arquivo				
Dimensões					
Termo de autoridade	Termo inserido				
Colaborador	Quem insere o registo de autoridade				
Tempo	Data de inserção do registo de autoridade				
Medidas					
Registo	Número de registos de autoridade novos inseridos na Biblioteca, Museu ou arquivo				

Estrela	Autoridades alteradas	Versão	1.2	Data	05/04/2013
Granularidade	Registos de autoridade alterados no sistema da Biblioteca, Museu ou arquivo				
Dimensões					
Termo de autoridade	Termo alterado				
Colaborador	Quem altera o registo de autoridade				
Tempo	Data de alteração do registo de autoridade				
Medidas					
Registo de autoridade	Número de registos de autoridade alterados na Biblioteca, Museu ou arquivo				

Estrela	Indexação	Versão	1.0	Data	05/05/2013
Granularidade	Documento indexado				
Dimensões					
Documento	Documento				
Grupo de assuntos	Tabela de ponte para o conjunto de assuntos atribuídos a um documento				
Medidas					

Estrela	Tratamento técnico	Versão	1.0	Data	05/05/2013
Granularidade	Tarefa executada em um documento				
Dimensões					
Documento	Documento				
Colaborador	Colaborador do serviço autor da tarefa				
Tarefa (DD)	As diversas fases pelas quais passa o documento no momento do seu tratamento técnico				
Tempo	Tempo				
Medidas					
Tempo	Tempo que uma tarefa demora a ser executada num documento				

Estrela	Controlo de qualidade	Versão	1.0	Data	07/02/2013
Granularidade	Inconformidade detetada em um documento				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Quem altera o registo de autoridade				
Tarefa	As diversas fases pelas quais passa o documento no momento do seu tratamento técnico				
Tempo	Tempo				
Medidas					
Quantidade	Número de não conformidades				

Estrela	Outras colaborações	Versão	1.1	Data	08/02/2013
Granularidade	Colaboração de uma pessoa em um evento				
Dimensões					
Colaborador	Colaborador da Biblioteca				
Evento	Evento				
Tempo	Tempo				
Medidas					
Colaborações	Número de colaborações dos membros da equipa em outras atividades				

F.2 Factos dos *data marts* da equipa EAD

Estrela	Inscrições	Versão	2.0	Data	13/05/2013
Granularidade	Inscrição de um utilizador na Biblioteca				
Dimensões					
Utilizador	Pessoa que se inscreve na Biblioteca				
Estatuto (DD)	Estatuto do utilizador				
MesRef	Mês em que contagem foi efetuada				
Validade	Data de validade do registo do utilizador, quando foi efetuada a contagem				
Medidas					
Quantidade	Utilizadores inscritos e com conta ativa				

Estrela	Multas	Versão	2.1	Data	25/05/2013
Granularidade	Multa cobrada a utilizador, resultante da sua atividade... multas, inscrições, etc...				
Dimensões					
Documento	Documento				
Colaborador	Colaborador que cobra a multa				
Utilizador	A quem é cobrada a multa				
Estado (DD)	Estado da transação (cobrada, não cobrada)				
Tipo (DD)	Tipo de transação (multas, inscrições)				
Recibo (DD)	Número do recibo				
Tempo	Data em que a transação foi lançada				
Tempo	Data de pagamento da multa				
Medidas					
Valor	Valor em euros de multas cobradas a utilizadores				

Estrela	Circulação	Versão	2.0	Data	17/05/2013
Granularidade	Circulação ou Empréstimo de um exemplar de um documento a um utilizador, em determinada data, por determinado colaborador				
Dimensões					
Documento	Exemplar de documento associado ao evento.				
Colaborador	Colaborador responsável pelo acontecimento, quando aplicável.				
Utilizador	Quem despoleta o evento de circulação ou empréstimo.				
Curso	No caso de o utilizador ser um estudante, representa o Curso ao qual o utilizador está associado no momento da realização do evento				
UO	Caso o utilizador seja um colaborador, representa a sua unidade orgânica				
Tipo material (DD)	Tipo de material envolvido na transação (LIVRO, TESE, VIDEO, etc...)				
Subbib (DD)	Sub-biblioteca à qual o documento se encontra associado				
Estatuto exemplar (DD)	Estatuto que o exemplar detinha no momento no tempo em que o evento foi realizado				
Estatuto utilizador (DD)	Estatuto que o utilizador detinha no momento no tempo em que o evento foi realizado				
Tipo utilizador (DD)	Classificação que o utilizador detinha quando o evento ocorreu (Aluno, Docente, Funcionário, investigador, etc...)				
Tipo evento (DD)	Tipo de evento registado no cubo (51-Empréstimo;61-Devolução;73,74-Reserva;62-64-Renovação;82-Consulta presencial)				
Tempo	Tempo				
Medidas					
Quantidade	Número de ocorrências de determinado tipo de evento				

Estrela	Obras extraviadas	Versão	2.0	Data	17/05/2013
Granularidade	Exemplar de documento cujo paradeiro é desconhecido, quer por perda resultante de empréstimo a utilizador, quer por não estar localizado onde seria devido encontrar-se, dentro das instalações da Biblioteca.				
Dimensões					
Documento	Documento				
Estado Processamento (DD)	Estado de processamento atribuído ao exemplar. Pode assumir um de três valores: IN - Indisponível; DS - Desaparecido; EX - Extraviado				
Colaborador	Utilizador com empréstimo atrasado				
Tempo	Tempo				
Medidas					
Quantidade	Número de exemplares de documento extraviados				

Estrela	Consultas acompanhadas ao arquivo	Versão	2	Data	25/05/2013
Granularidade	Minutos gastos em consultas ao arquivo, efetuada por um utilizador ou entidade, em um determinado mês				
Dimensões					
Colaborador	Colaborador responsável pela operação				
Utilizador	Utilizador que pede para consultar processo do Arquivo				
UO	Afiliação do utilizador que consulta				
Tipo consulta (DD)	Natureza da consulta (administrativa, de investigação, de manutenção...)				
Tipo contato (DD)	Natureza do contato (presencial, telefone, email...)				
Tempo	Mês em que as consultas foram efetuadas				
Medidas					
Duração	Duração em minutos da consulta				

Estrela	Consultas desacompanhadas ao arquivo	Versão	2.0	Data	25/05/2013
Granularidade	Numero de consultas ao arquivo efetuada por uma entidade, por mês				
Dimensões					
UO	Afiliação do utilizador que consulta				
Tipo consulta (DD)	Natureza da consulta (administrativa, de investigação, de manutenção...)				
Tempo	Mês em que as consultas foram efetuadas				
Medidas					
Consultas	Número de consultas sem acompanhamento efetuadas por um departamento/serviço, por mês				

Estrela	Requisições ao arquivo	Versão	2.0	Data	25/05/2013
Granularidade	Número de requisições e os documentos envolvidos, efetuados por determinada entidade, por mês				
Dimensões					
UO	Entidade requisitante				
Tempo	Tempo				
Medidas					
Requisições	Número de requisições				
Documentos	Número de documentos				

Estrela	Pedidos de apoio de 1.ª linha	Versão	2.0	Data	25/05/2013
Granularidade	Pedido de auxílio de um utilizador a um colaborador				
Dimensões					
Colaborador	Colaborador responsável pelo apoio ao utilizador				
Utilizador	Utilizador que faz pedido de apoio				
Categoria	Caracterização do tipo de pedido				
Piso (DD)	Piso da Biblioteca em que o pedido é efetuado				
Pedido (DD)	Registo em texto livre do pedido de apoio efetuado				
Estado (DD)	Estado de resolução do pedido				
Tipo utilizador (DD)	Caracterização do perfil do utilizador que faz o pedido de apoio				
Tempo	Tempo				
Medidas					
Pedidos	Número de pedidos de apoio de primeira linha recebidos				

Estrela	Ocupação pisos	Versão	2.0	Data	25/05/2013
Granularidade	Ocupação nos pisos da Biblioteca, medidos diariamente de duas em duas horas, por piso				
Dimensões					
Piso (DD)	Piso da biblioteca				
Ronda (DD)	Identificação da passagem do segurança				
Tempo	Tempo				
Medidas					
PP	Número de utilizadores no piso				
LO	Número de lugares de leitura ocupados				
CO	Número de computadores ocupados				

Estrela	Fotocopiadoras	Versão	2.0	Data	25/05/2013
Granularidade	Número de fotocópias tiradas nas máquinas dos pisos, medido mensalmente				
Dimensões					
Piso (DD)	Piso da biblioteca				
Tempo	Tempo				
Medidas					
LCF	Leitura do contador de fotocópias				
F	Número de fotocópias tiradas nas máquinas dos pisos, medido mensalmente				

Estrela	Reserva de gabinetes	Versão	2.0	Data	25/05/2013
Granularidade	Pedido de reserva de um gabinete por um utilizador				
Dimensões					
Colaborador	Colaborador responsável por empréstimo				
Utilizador	Utilizador que requisita equipamento				
Estado (DD)	Estado do pedido				
Gabinete (DD)	Gabinete alvo do pedido de reserva				
Período Reserva (DD)	Período para o qual o gabinete é pedido				
Tempo	Data em que o pedido foi colocado				
Tempo	Data em que a ocupação do gabinete é consumada				
Medidas					
Reservas satisfeitas	Número de pedidos de reserva aprovados (Q=1 => reserva aprovada; Q=0 => reserva não aprovada)				

Estrela	Formação	Versão	2.0	Data	27/05/2013
Granularidade	Sessão de formação realizada				
Dimensões					
Colaborador	Colaborador responsável pela formação				
Evento	Sessão de formação realizada				
UC	Unidade Curricular na qual a sessão de formação está inserida				
Curso	Curso para o qual a sessão de formação foi realizada				
Tempo	Tempo				
Medidas					
Alunos	Número de alunos que participaram				
Duração	Duração de cada sessão de formação				
Sessões	Número de sessões de formação realizadas				

Estrela	Pedidos de apoio de 2.ª linha	Versão	2.0	Data	27/05/2013
Granularidade	Pedido de auxílio de um utilizador a um colaborador				
Dimensões					
Colaborador	Colaborador responsável pelo apoio ao utilizador				
Utilizador	Utilizador que faz pedido de apoio				
Pedido (DD)	Registo em texto livre do pedido de apoio efetuado				
Tipo contato (DD)	Forma como o serviço foi abordado [email presencial telefone]				
Tipo utilizador (DD)	Caracterização do perfil do utilizador que faz o pedido de apoio				
Tempo	Tempo				
Medidas					
Pedidos	Número de pedidos de apoio de segunda linha recebidos				

Estrela	Participação	Versão	2.0	Data	26/05/2013
Granularidade	Utilizador que participa em sessão de formação				
Dimensões					
Utilizador	Utilizador que participa em sessão de formação				
Colaborador	Colaborador responsável pela formação				
Curso	Curso do utilizador que frequenta a sessão de formação (quando aplicável)				
Sessão formação (DD)	Sessão de formação realizada				
Tipo utilizador (DD)	Caracterização do perfil do participante na ação de formação				
Presença (DD)	Indicação se utilizador compareceu, faltou, cancelou ou adiou a formação				
Certificado (DD)	Indicação se sessão dá acesso a certificado de formação				
Tempo	Tempo				
Medidas					
Presença	Número de formandos que participaram em sessões de formação				
Duração	Duração da sessão da sessão de formação (horas)				

Estrela	Utilizadores externos	Versão	1.0	Data	12/02/2013
Granularidade	Inscrição de um utilizador externo na Biblioteca				
Dimensões					
Colaborador	Colaborador responsável pela inscrição				
Utilizador	Utilizador externo inscrito				
Tempo	Tempo				
Medidas					
Inscrições	Número de colaboradores externos inscritos				

Estrela	Sessões	Versão	1.0	Data	12/02/2013
Granularidade	Sessão web registada nos fornecedores de recursos de informação licenciados pela FEUP				
Dimensões					
Recurso de pesquisa	Fornecedores de recursos de informação licenciados pela FEUP				
Tempo	Tempo				
Medidas					
Sessões	Número de sessões estabelecidas a recursos de pesquisa contratados pela FEUP				

Estrela	Pesquisas	Versão	2.0	Data	28/05/2013
Granularidade	Pesquisa feita por um utilizador em um recurso de pesquisa				
Dimensões					
Recurso de pesquisa	Fornecedores de recursos de informação licenciados pela FEUP				
Estado pesquisa (DD)	Resultado da operação de pesquisa [Concluída Suspensa Erro Cancelada ...]				
Tempo	Tempo				
Medidas					
Pesquisas	Número de pesquisas efetuadas em recursos de informação				

Estrela	Visualizações	Versão	1.0	Data	12/02/2013
Granularidade	Visualização de um documento em formato eletrónico, disponível num fornecedor de recursos de informação.				
Dimensões					
Documento	Documento eletrónico visualizado				
Recurso de pesquisa	Fornecedor de conteúdos através do qual o documento foi visualizado				
Tempo	Tempo				
Medidas					
Visualizações	Número de visualizações				

Estrela	TT EAD	Versão	1.0	Data	12/02/2013
Granularidade	Trouble ticket colocado à Equipa EAD				
Dimensões					
Colaborador	Colaborador responsável pelo TT				
Utilizador	Utilizador que coloca o TT				
UO	Afiliação do utilizador				
Tempo	Tempo				
Medidas					
Colocados	Número de TT colocados				
Respondidos	Número de TT respondidos				

Estrela	Entradas na Biblioteca.	Versão	1.0	Data	12/02/2013
Granularidade	N.º de pessoas que entram na Biblioteca, por dia, medida na antena da entrada da Biblioteca.				
Dimensões					
Tempo	Tempo				
Medidas					
Entradas	Número de entradas				

Estrela	Receção Material	Versão	1.0	Data	12/02/2013
Granularidade	Exemplar rececionado no piso zero da Biblioteca				
Dimensões					
Documento	Exemplar do documento rececionado				
Colaborador	Colaborador que receciona o material				
Utilizador	Utilizador que entrega o material				
UO	Afiliação do utilizador				
Tempo	Tempo				
Medidas					
Receção	Número de documentos rececionados				

Estrela	Extravios	Versão	1.0	Data	12/02/2013
Granularidade	Exemplar de documento extraviado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador que assinala o extravio				
Tempo	Tempo				
Medidas					
Extravios	Número de documentos extraviados				

Estrela	Problemas comportamentais	Versão	1.0	Data	12/02/2013
Granularidade	Registo de um incidente comportamental de um utilizador na Biblioteca				
Dimensões					
Utilizador	Utilizador responsável pelo incidente				
Tempo	Tempo				
Medidas					
Incidente	Número de incidentes comportamentais registados				

Estrela	Danos	Versão	1.0	Data	12/02/2013
Granularidade	Exemplar de documento danificado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador que assinala o documento danificado				
Tempo	Tempo				
Medidas					
Danos	Número de documentos danificados				

Estrela	Sanções	Versão	1.0	Data	12/02/2013
Granularidade	Sansão aplicada a um utilizador				
Dimensões					
Utilizador	Utilizador sancionado				
Tempo	Tempo				
Medidas					
Sanções	Número de sanções aplicadas a utilizadores				

F.3 Factos dos data marts da equipa EMAP

Estrela	Oferta Documentos	Versão	1.0	Data	13/02/2013
Granularidade	Oferta de um exemplar de uma obra à Biblioteca				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Utilizador	Utilizador que oferece obra				
Tempo	Tempo				
Medidas					
Ofertas	Número de obras oferecidas				

Estrela	Ofertas Museu	Versão	1.0	Data	13/02/2013
Granularidade	Oferta de um objeto ao Museu				
Dimensões					
Documento	Documento				
Utilizador	Utilizador que oferece documento				
Tempo	Tempo				
Medidas					
Ofertas	Número de ofertas ao museu				

Estrela	Ocupação depósitos	Versão	1.0	Data	13/02/2013
Granularidade	Ocupação de um exemplar de um documento ou processo de arquivo, medido em metros lineares				
Dimensões					
Documento	Documento				
UO	Afiliação do processo/documento				
Sala	Sítio onde está localizado o depósito				
Estante depósito	Identificação da estante do depósito				
Tempo	Tempo				
Medidas					
Ocupação	Espaço ocupado por processo/documento				

Estrela	Incorporações na Biblioteca	Versão	1.0	Data	13/02/2013
Granularidade	Exemplar de documento incorporado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela receção do processo				
Utilizador	Utilizador responsável pela entrega do processo no Arquivo				
Tempo	Tempo				
Medidas					
Incorporações	Número de exemplares de documento incorporado				

Estrela	Transferências no Arquivo	Versão	1.0	Data	29/10/2012
Granularidade	Transferência de um processo				
Dimensões					
Documento	Processo				
Colaborador	Responsável pela transferência				
Departamento/Serviço	Origem				
Tempo	Data da transferência				
Medidas					
Processo	Número de processos transferidos				

Estrela	Alienações no Arquivo	Versão	1.0	Data	13/02/2013
Granularidade	Alienação de um processo do arquivo				
Dimensões					
Documento	Processo				
Colaborador	Responsável pela alienação				
UO	Afiliação do processo				
Sala	Sítio onde está localizado o processo				
Estante depósito	Identificação da estante do depósito				
Tempo	Tempo				
Medidas					
Alienações	Número de alienações no arquivo				

Estrela	Alienações na Biblioteca	Versão	1.0	Data	13/02/2013
Granularidade	Alienação de um exemplar de um documento				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela alienação do documento				
Tempo	Tempo				
Medidas					
Alienações	Número de alienações na Biblioteca				

Estrela	Instalação Documentos	Versão	1.0	Data	13/02/2013
Granularidade	Instalação/reinstalação de um exemplar de um documento em depósito ou reserva				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela Instalação/reinstalação do documento				
Tempo	Tempo				
Medidas					
Instalações	Número de instalações/reinstalações de documentos em depósito ou reserva;				

Estrela	Instalação Museu	Versão	1.0	Data	13/02/2013
Granularidade	Instalação/reinstalação de um objeto museológico, em depósito ou reserva;				
Dimensões					
Documento	Documento				
Colaborador	Colaborador responsável pela Instalação/reinstalação do documento				
Tempo	Tempo				
Medidas					
Instalações	Número de instalações/reinstalações de objetos museológicos, em depósito ou reserva;				

Estrela	Monitorização Ambiental	Versão	1.0	Data	14/02/2013
Granularidade	Humidade, Temperatura e Luminosidade medida periodicamente nos espaços de depósito e reserva				
Dimensões					
Sala	Sítio monitorizado				
Humidade	Humidade				
Temperatura	Temperatura				
Luminosidade	Luminosidade				
Tempo	Tempo				
Medidas					
Humidade	Percentagem de humidade				
Temperatura	Graus centígrados				
Luminosidade	Lúmen medidos				

Estrela	Higienização Documentos	Versão	1.0	Data	14/02/2013
Granularidade	Exemplar de documento higienizado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela intervenção				
Tempo	Tempo				
Medidas					
Exemplares	Número de exemplares higienizados				

Estrela	Higienização Museu	Versão	1.0	Data	14/02/2013
Granularidade	Documentos de museu higienizados				
Dimensões					
Documento	Documento				
Colaborador	Colaborador responsável pela intervenção				
Tempo	Tempo				
Medidas					
Objetos	Objetos do museu higienizados				

Estrela	Digitalização	Versão	1.0	Data	14/02/2013
Granularidade	Exemplar de documento digitalizado para preservação				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela intervenção				
Tempo	Tempo				
Medidas					
Digitalizações	Número de digitalizações realizadas				

Estrela	Encadernação	Versão	1.0	Data	14/02/2013
Granularidade	Exemplar de documento encadernado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador responsável pela intervenção				
Tempo	Tempo				
Medidas					
Encadernações	Número de exemplares de documentos encadernados				

Estrela	Produção	Versão	1.0	Data	14/02/2013
Granularidade	Documento em formato eletrónico produzido				
Dimensões					
Documento	Documento				
Colaborador	Colaborador (es) responsável (eis) pela produção do documento				
Tempo	Tempo				
Medidas					
Documentos	Número de documentos produzidos				

Estrela	Depósito	Versão	1.0	Data	14/02/2013
Granularidade	Documento depositado no repositório institucional				
Dimensões					
Documento	Documento				
Colaborador	Colaborador responsável pelo depósito				
Tempo	Tempo				
Medidas					
Depósitos	N.º de documentos depositados				

Estrela	Disponibilização	Versão	1.0	Data	14/02/2013
Granularidade	Documento disponibilizado				
Dimensões					
Documento	Documento				
Tempo	Tempo				
Medidas					
Tempo	Número de dias necessários para disponibilizar os documentos				

Estrela	Ebooks	Versão	1.0	Data	14/02/2013
Granularidade	Tarefas para a produção de um E-book				
Dimensões					
Documento	Documento do tipo <i>Ebook</i>				
Colaborador	Colaborador responsável pela realização da tarefa				
Tarefa	Parte constituinte do processo de produção de um E-book				
Tempo	Tempo				
Medidas					
Tempo	Tempo de produção de um E-book				

Estrela	Satisfação <i>Ebooks</i>	Versão	1.0	Data	14/02/2013
Granularidade	<i>Ebook</i> produzido				
Dimensões					
Documento	Documento do tipo <i>Ebook</i>				
Utilizador	Cliente e/ou responsável da coleção				
Tempo	Tempo				
Medidas					
Satisfação	Grau de satisfação do cliente e do responsável da coleção quanto ao <i>ebook</i> produzido				

F.4 Factos dos *data marts* da equipa ESE

Estrela	Desenvolvimento Projetos	Versão	1.0	Data	14/02/2013
Granularidade	Tarefa do projeto para o desenvolvimento de funcionalidades				
Dimensões					
Funcionalidade	Funcionalidade em desenvolvimento				
Utilizador	Utilizador responsável pela tarefa				
Projeto	Projeto de desenvolvimento				
Tarefa	Parte constituinte do projeto de desenvolvimento				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto no desenvolvimento de funcionalidades				

Estrela	Satisfação Sítios Web	Versão	1.0	Data	14/02/2013
Granularidade	Sítio temático desenvolvido				
Dimensões					
Sítio temático	Sítio Web desenvolvido para um evento				
Utilizador	Cliente que pediu desenvolvimento de sítio temático				
Evento	Evento				
Tempo	Tempo				
Medidas					
Satisfação	Grau de satisfação do cliente pelo Site Temático desenvolvido				

Estrela	Desenvolvimento <i>Web design</i>	Versão	1.0	Data	14/02/2013
Granularidade	Tarefa do projeto para o desenvolvimento do Site temático, Serviço <i>WEB</i> ou Material de Comunicação				
Dimensões					
Colaborador	Colaborador responsável pela tarefa				
Projeto	Projeto de desenvolvimento				
Tarefa	Parte constituinte do projeto (pode ser do tipo: Site Temático, Serviço <i>WEB</i> , Material de comunicação, <i>Banners</i> , etc.)				
Tempo	Tempo				
Medidas					
Tempo	Tempo de desenvolvimento dos serviços e materiais de comunicação				

Estrela	Disponibilidade	Versão	1.0	Data	14/02/2013
Granularidade	Aplicação de suporte				
Dimensões					
Aplicação de suporte	Aplicações informáticas do serviço				
Tempo	Tempo				
Medidas					
Tempo	Taxa de disponibilidade dos serviços				

Estrela	Reclamações ESE	Versão	1.0	Data	14/02/2013
Granularidade	Reclamação realizada por um utilizador relacionada com um serviço web, uma aplicação de suporte ou funcionalidade				
Dimensões					
Reclamação	Reclamação sobre o mal funcionamento de um serviço web, uma aplicação de suporte ou funcionalidade				
Utilizador	Utilizador que faz reclamação				
Tempo	Tempo				
Medidas					
Reclamações	Número de reclamações				

Estrela	Manutenção	Versão	1.0	Data	14/02/2013
Granularidade	Intervenção realizada numa aplicação de suporte				
Dimensões					
Aplicação de suporte	Aplicação intervencionada				
Colaboração	Utilizador responsável pela manutenção				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto na manutenção da infraestrutura				

Estrela	Desenvolvimento aplicações	Versão	1.0	Data	14/02/2013
Granularidade	Desenvolvimento de uma funcionalidade numa aplicação de suporte				
Dimensões					
Aplicação de suporte	Aplicação alvo do desenvolvimento				
Funcionalidade	Funcionalidade desenvolvida				
Colaborador	Utilizador responsável pelo desenvolvimento				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto no desenvolvimento de funcionalidades nas aplicações de suporte				

Estrela	Satisfação desenvolvimentos	Versão	1.0	Data	14/02/2013
Granularidade	Funcionalidade desenvolvida				
Dimensões					
Funcionalidade	Funcionalidade desenvolvida				
Utilizador	Cliente				
Tempo	Tempo				
Medidas					
Satisfação	Grau de satisfação do cliente face às soluções desenvolvidas				

F.5 Factos dos *data marts* da equipa AGORA

Estrela	Conceção	Versão	1.0	Data	14/02/2013
Granularidade	Evento criado				
Dimensões					
Colaborador	Colaborador responsável pela conceção				
Projeto	Projeto do evento				
Evento	Evento				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto na conceção de projeto				

Estrela	Desenvolvimento AGORA	Versão	1.0	Data	14/02/2013
Granularidade	Atividade desenvolvida para a construção de um evento				
Dimensões					
Colaborador	Colaborador responsável pela tarefa				
Projeto	Projeto do evento				
Evento	Evento				
Tarefa	Parte constituinte do projeto do evento				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto no desenvolvimento das atividades				

Estrela	Interpretação	Versão	1.0	Data	14/02/2013
Granularidade	Interpretação de um documento				
Dimensões					
Documento	Documento alvo da interpretação				
Colaborador	Colaborador (es) responsável (eis) pela interpretação				
Evento	Evento no âmbito do qual o documento esta a ser interpretado				
Tarefa	Tarefa de interpretação				
Tempo	Tempo				
Medidas					
Tempo	Tempo gasto no trabalho de interpretação				

Estrela	Participação	Versão	1.0	Data	14/02/2013
Granularidade	Participação de uma pessoa em um evento				
Dimensões					
Evento	Evento realizado				
Tempo	Tempo				
Medidas					
Participação	Índices de participação nos eventos				

Estrela	Satisfação AGORA	Versão	1.0	Data	14/02/2013
Granularidade	Satisfação de uma pessoa relativamente a um evento				
Dimensões					
Utilizador	Participante em um evento				
Evento	Evento realizado				
Tempo	Tempo				
Medidas					
Satisfação	Grau de satisfação do participante de um evento				

Estrela	Relações	Versão	1.0	Data	14/02/2013
Granularidade	Relação entre SDI e um Departamento, Serviço ou Instituto resultante da realização de um evento				
Dimensões					
UO	Parceiro na realização em um evento				
Evento	Evento realizado				
Tempo	Tempo				
Medidas					
Relação	Relações estabelecidas entre SDI/Departamentos e Serviços FEUP (ou seja, mapear as relações de forma a assegurar a representatividade de todos os Departamentos/Serviços nas atividades culturais)				

Estrela	Impacto	Versão	1.0	Data	14/02/2013
Granularidade	Impacto de um evento no número de empréstimos de um documento				
Dimensões					
Documento	Documento alvo da avaliação de impacto				
Classificação	Classificação CDU atribuída aos documentos				
Evento	Evento realizado				
Tempo	Tempo				
Medidas					
Impacto	Impacto das atividades sobre movimento dos acervos				

F.6 Factos dos *data marts* da equipa ESA

Estrela	Visitas	Versão	1.0	Data	15/02/2013
Granularidade	Visita oficial de uma ou mais pessoas à Biblioteca				
Dimensões					
Colaborador	Pessoa (s) que recebem e fazem a visita guiada				
Utilizador	Identificação do grupo de pessoas que visita a Biblioteca				
Tempo	Tempo				
Medidas					
Pessoas	Número de pessoas que compõem a delegação				

Estrela	Execução PA	Versão	1.0	Data	15/02/2013
Granularidade	Tarefa de um plano de atividades				
Dimensões					
Plano de atividade	Plano de atividade do SDI				
Tarefa	Parte constituinte de um plano de atividades				
Tempo	Tempo				
Medidas					
Percentagem	Execução do plano de atividades				

Estrela	Instruções de trabalho novas	Versão	1.0	Data	15/02/2013
Granularidade	Instrução de trabalho criada por uma pessoa				
Dimensões					
Colaborador	Colaborador que cria nova instrução de trabalho				
Instrução de trabalho	Documento que define regras e condutas relacionadas com determinado processo de trabalho				
Tempo	Tempo				
Medidas					
Instruções	Número de instruções de trabalho novas				

Estrela	Instrução de trabalho corrigidas	Versão	1.0	Data	15/02/2013
Granularidade	Instrução de trabalho corrigida por uma pessoa				
Dimensões					
Colaborador	Colaborador que corrige instrução de trabalho				
Instrução de trabalho	Documento que define regras e condutas relacionadas com determinado processo de trabalho				
Tempo	Tempo				
Medidas					
Instruções	Número de instruções de trabalho corrigidas				

Estrela	Projetos cooperação	Versão	1.0	Data	15/02/2013
Granularidade	Projeto de cooperação e integração a nível nacional ou internacional, no qual o SDI está envolvido				
Dimensões					
Colaborador	Colaboradores envolvidos no projeto				
Projeto cooperação	Projeto de cooperação e integração a nível nacional ou internacional, no qual o SDI está envolvido				
Tempo	Tempo				
Medidas					
Projetos	Número de projeto de cooperação e integração a nível nacional ou internacional, no qual o SDI está envolvido				

Estrela	Reclamações ESA	Versão	1.0	Data	15/02/2013
Granularidade	Reclamação efetuada por um utilizador				
Dimensões					
Reclamação	Reclamação efetuada pelo utilizador				
Utilizador	Utilizador que faz a reclamação				
Tempo	Tempo				
Medidas					
Reclamação	Número de reclamações dos utentes				

Estrela	Condições ambientais	Versão	1.0	Data	15/02/2013
Granularidade	Temperatura e humidade medida nos diversos locais que compõem a Biblioteca				
Dimensões					
Sala	Piso e/ou sala				
Temperatura	Temperatura registada				
Humidade	Humidade registada				
Tempo	Tempo				
Medidas					

Estrela	Gestão espaços	Versão	1.0	Data	15/02/2013
Granularidade	Ocorrência relacionada com o estado de conservação e funcionamento de determinando espaço.				
Dimensões					
Colaborador	Colaborador que regista incidente				
Sala	Local onde ocorreu incidente				
Incidente	Ocorrência relacionada com o estado de conservação e funcionamento de determinando espaço				
Tempo	Tempo				
Medidas					
Incidente	Número de incidentes relacionados com o estado de conservação e funcionamento de determinando espaço				

Estrela	Gestão equipamentos	Versão	1.0	Data	15/02/2013
Granularidade	Ocorrência relacionada com o estado de conservação e funcionamento de determinando equipamento.				
Dimensões					
Colaborador	Colaborador que regista incidente				
Sala	Local onde ocorreu incidente				
Equipamento	Equipamento afetado				
Incidente	Ocorrência relacionada com o estado de conservação e funcionamento de determinando equipamento				
Tempo	Tempo				
Medidas					
Incidente	Número de incidentes relacionados com o estado de conservação e funcionamento de determinando equipamento				

Estrela	Inventário	Versão	1.0	Data	15/02/2013
Granularidade	Número de cadeiras, mesas, computadores. Bens confirmados/Não confirmados.				
Dimensões					
Colaborador	Colaborador envolvido no processo				
Sala	Local do equipamento				
Equipamento	Material inventariado				
Tempo	Tempo				
Medidas					
Quantidade	Número de itens do material inventariado				

Estrela	Movimentos	Versão	1.0	Data	15/02/2013
Granularidade	Movimento financeiro relacionado com a transação de um item				
Dimensões					
Utilizador	Utilizador relacionado com a transação				
Orçamento	Orçamento onde a despesa será cabimentada				
Rubrica	Rubricas previamente definidas, por ex. bibliografia, material escritório, formação, <i>ebooks</i> , etc...				
CCO	Centro de Custos Operacional				
Produto	Produto transacionado				
PAD	Pedido de Autorização de Despesa				
Fornecedor	Fornecedor do produto transacionado				
Tempo	Tempo				
Medidas					
Custo	Custos efetuados. Total gasto, recebido e saldo.				

Estrela	Bibliografia UC	Versão	1.1	Data	20/02/2013
Granularidade	Título da obra da Bibliografia de uma UC, lecionada em um curso, numa determinada ocorrência				
Dimensões					
Bibliografia UC	Obra que faz parte da Bibliografia de uma UC				
Unidade Curricular	Unidade Curricular lecionada				
Curso	Curso				
Ocorrência	Ano e semestre em que UC é lecionada				
UO	Departamento onde a UC está afiliada				
Tempo	Tempo				
Medidas					
Existência	Bibliografia das unidades curriculares que existem na biblioteca ou não				

Estrela	Comparativo	Versão	1.0	Data	15/02/2013
Granularidade	Artigo adquirido				
Dimensões					
Produto	Artigo adquirido				
Fornecedor	Fornecedor do artigo adquirido				
Tempo	Tempo				
Medidas					
Quantidade	Quantidade de artigos adquiridos				
Preço	Preço unitário				

Estrela	Reposições	Versão	1.0	Data	15/02/2013
Granularidade	Exemplar de documento extraviado				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador envolvido na reposição do extravio				
Utilizador	Utilizador envolvido no extravio				
Tempo	Tempo				
Medidas					
Reposições	N.º de obras extraviadas repostas				

Estrela	Aquisições	Versão	1.0	Data	15/02/2013
Granularidade	Exemplar de um documento adquirido				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Colaborador	Colaborador envolvido na aquisição				
Utilizador	Utilizador envolvido na aquisição				
Orçamento	Orçamento onde a despesa será cabimentada				
CCO	Centro de Custos Operacional				
PAD	Pedido de Autorização de Despesa				
Fornecedor	Fornecedor do produto transacionado				
Tempo	Tempo				
Medidas					
Tempo	Tempo de aquisição de documentos				

Estrela	EIB	Versão	1.0	Data	15/02/2013
Granularidade	Exemplar de um documento transacionado por EIB				
Dimensões					
Documento	Documento				
Classificação	Classificação CDU atribuída aos documentos				
Indexação	Os assuntos abordados nos documentos				
Material EIB (DD)	Tipo de material (livro, norma, tese, etc.). Tipo de transação (cópia, empréstimo).				
Colaborador	Colaborador envolvido na aquisição				
Utilizador	Utilizador envolvido na aquisição				
CCO	Centro de Custos Operacional				
PAD	Pedido de Autorização de Despesa				
Fornecedor	Fornecedor do produto transacionado				
Tempo	Tempo				
Medidas					
Tempo	Tempo de fornecimento de documentos. Tempos mínimos, médios e máximos por tipo de material.				
Pedidos	Número de pedidos				

Estrela	Vendas	Versão	1.0	Data	01/05/2013
Granularidade	Exemplar de obra comercializado				
Dimensões					
Documento	Título comercializado				
Cliente	Cliente que compra				
Tempo	Tempo				
Medidas					
Número	Número de exemplares vendidos de uma obra, em uma determinada data				
Valor	Preço de venda de um exemplar, numa determinada data				

Estrela	Preparação	Versão	1.0	Data	01/05/2013
Granularidade	Título comercializado pela FEUP Edições				
Dimensões					
Documento	Título comercializado				
Tarefa	Tarefa a ser realizada no âmbito da preparação de uma obra				
Tempo inicia	Data de início da tarefa				
Tempo fim	Data de fim da tarefa				
Medidas					
Dias	Número de dias gastos para a realização de uma determinada tarefa de preparação				

Estrela	Produção	Versão	1.0	Data	01/05/2013
Granularidade	Título comercializado pela FEUP Edições				
Dimensões					
Documento	Título comercializado				
Tarefa	Tarefa a ser realizada no âmbito da produção de uma obra				
Tempo inicia	Data de início da tarefa				
Tempo fim	Data de fim da tarefa				
Medidas					
Dias	Número de dias gastos para a realização de uma determinada tarefa de produção				

Anexo G. Dicionário de dimensões

Nome	Descrição	SCD	Versão	1.4	Data	05/05/2013	Fontes
DOCUMENTO	Qualquer manifestação de um qualquer objeto elaborada com o fim de reproduzir ou representar uma pessoa, um facto, um dito ou um acontecimento.	Tipo 1	Hierarquia	Documento < Exemplar			
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	ALEPH; Digitool; Inarte; Gisa
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
EXEMPLAR_ID	Identificador de exemplar	Exemplar	BK	VARCHAR2			
CODIGO_BARRAS	Código de barras atribuído ao exemplar	Exemplar	UK	VARCHAR2	30		
ESTATUTO	Estatuto de exemplar	Exemplar		CHAR	2		
SUB_BIBLIOTECA	Sub-biblioteca atribuído ao exemplar	Exemplar		VARCHAR2	10		
TIPO_MATERIAL	Tipo de material do exemplar	Exemplar		VARCHAR2	5		
COLEÇÃO	Para a Bib é o piso onde o exemplar está guardado. Para o museu é a coleção propriamente dita	Exemplar		VARCHAR2	200		
COTA	Cota atribuída ao exemplar	Exemplar		VARCHAR2	100		
DESCRIÇÃO	Campo de texto livre para descrição do exemplar	Exemplar		VARCHAR2	255		
DATA_FASCÍCULO	Data de publicação de um fascículo, no caso de exemplares do tipo periódico	Exemplar		DATE			
DATA_RECECAO_ESPERADA	Data esperada para receção de um fasciculo	Exemplar		DATE			
DATA_RECECAO	Data de receção de um fascículo	Exemplar		DATE			
ENUMERACAO_A	Campo a para enumerar fasciculos de publicações periódicas	Exemplar		VARCHAR2	20		
ENUMERACAO_B	Campo b para enumerar fasciculos de publicações periódicas	Exemplar		VARCHAR2	20		
ENUMERACAO_C	Campo c para enumerar fasciculos de publicações periódicas	Exemplar		VARCHAR2	20		
CRONOLOGIA_I	Cronologia i para fascículos de publicações periódicas	Exemplar		VARCHAR2	20		
CRONOLOGIA_J	Cronologia j para fascículos de publicações periódicas	Exemplar		VARCHAR2	20		
DOC_ID	Identificador unívoco de documento	Documento	BK	VARCHAR2			
TITULO	Título	Documento		VARCHAR2	600		
ANO	Ano de publicação	Documento		NUMBER	4	0	
LINGUA	Língua de publicação	Documento		VARCHAR2	20		
PAIS	País de publicação	Documento		VARCHAR2	20		
EDITOR	Editor da obra	Documento		VARCHAR2	255		
ISSN_ISBN	Identificador ISBN ou ISSN do documento	Documento		VARCHAR2	100		
AUTOR	Chave para tabela de ponte que faz a ligação ao grupo de autores do documento	Documento		NUMBER	6	0	
CLASSE	Chave que faz ligação a dimensão que representa a classe do documento (CDU,ISAD,...)	Documento		VARCHAR2	15		
CLASSIFICAÇÃO	Classificação atribuída ao documento	Documento		VARCHAR2	500		
NATUREZA	Biblioteca/Arquivo/Museu/Doc.Elect./Ebook....	Documento		VARCHAR2	20		
NIVEL_UNIDADE_DESCRIÇÃO	Nível da unidade de descrição	Documento		VARCHAR2	20		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.6	<i>Data</i>	10/03/2013	<i>Fontes</i>
COLABORADOR	Pessoas que trabalham na FEUP e interagem com os sistemas, ficando aí registada a sua atividade	Tipo 2	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	ALEPH; Digitool; Metalib; SFX; SIGARRA; Inarte; GISA; Moddle; Sharepoint; Fontes <i>ad hoc</i>
DIMENSION_KEY	Chave da dimensão	Colaborador	PK	NUMBER			
NUMERO	Número mecanográfico do colaborador	Colaborador	BK	VARCHAR	10		
NOME	Nome do Colaborador	Colaborador		VARCHAR	100		
EMAIL	Endereço de correio eletrónico do colaborador	Colaborador		VARCHAR	50		
SEXO	Masculino/Feminino	Colaborador		CHAR	1		
CATEGORIA	Categoria profissional do colaborador	Colaborador		VARCHAR	50		
CARREIRA	Carreira profissional do colaborador	Colaborador		VARCHAR	50		
LOGIN	Login do colaborador	Colaborador		VARCHAR	50		
SERVIÇO	Departamento ou Serviço no qual o colaborador trabalha	Colaborador		VARCHAR	50		
DATA_EFETIVA	Data a partir do qual o registo é válido	Colaborador		DATE			
DATA_EXPIRACAO	Data de fim de validade do registo	Colaborador		DATE			

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	12/05/2013	<i>Fontes</i>
UTILIZADOR	São todos os potenciais utilizadores dos serviços oferecidos pelo SDI (ex. pessoas da FEUP e outras)	Tipo1	<i>Hierarquia</i>	Função<Pessoa			
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; <i>Aleph</i> ; <i>Digitool</i> ; <i>Metalib</i> ; Inarte; GISA; Fontes <i>ad hoc</i>
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
FUNCAOID	Código da função	Funcao	BK	VARCHAR	7		
FUNCAO	Função desempenhado pela pessoa	Funcao		VARCHAR	20		
IDUSER	Código principal do utilizador	Pessoa	BK	VARCHAR	25		
NOME	Nome do utilizador	Pessoa		VARCHAR	500		
LOGIN	Login	Pessoa		VARCHAR	50		
SEXO	Sexo do utilizador (M/F)	Pessoa		VARCHAR	1		
NASCIMENTO	Data de nascimento do Utilizador	Pessoa		DATE			
NATURALIDADE	Naturalidade do utilizador	Pessoa		VARCHAR	255		
EMAIL	<i>Email</i> do utilizador	Pessoa		VARCHAR	100		
TELEFONE	Telefone do utilizador	Pessoa		VARCHAR	50		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
TERMO_AUTORIDADE	É qualquer registo de um termo inserido numa base de autoridades (ex. nomes de autores, assuntos, títulos uniformes, etc...)	Tipo 1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	ALEPH; GISA; Inarte
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
IDAUT	Id do Registo de autoridade	Termo	BK	VARCHAR2			
TERMO	Termo controlado	Termo		VARCHAR2			
TIPO	Tipo de termo (autor, assunto, título, etc)	Termo		VARCHAR2			
ORIGEM	Origem do termo (Sistema/Serviço)	Termo		VARCHAR2			

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	17/05/2013	<i>Fontes</i>
CURSO	Um conjunto Unidades Curriculares estruturadas, que certificam competências e conferem um grau académico, depois frequentadas com sucesso.	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA
ID	Chave da dimensão		PK	NUMBER			
SIGLA	SIGLA do curso no Sigarra	Default	BK	VARCHAR2	10		
CODIGO	Código oficial do curso no DGES	Default		VARCHAR2	9		
NOME	Designação do curso	Default		VARCHAR2	100		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	17/05/2013	<i>Fontes</i>
UO	Departamentos e Serviços que fazem parte da estrutura orgânica da FEUP e ainda os Institutos de interface relacionados com a FEUP	Tipo1	<i>Hierarquia</i>	UO>UNIDADE>SUBUNIDADE			
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; GISA
ID	Chave da dimensão		PK	NUMBER			
UOID	ID da Unidade Orgânica (SIGLA)	UO	BK	VARCHAR2	20		
UO	Designação da UO	UO		VARCHAR2	200		
UNIDADEID	Sigla da unidade	UNIDADE	BK	VARCHAR2	20		
UNIDADE	Designação da unidade	UNIDADE		VARCHAR2	200		
TIPO_UNIDADE	Tipo de unidade, departamento, serviço, etc...	UNIDADE		VARCHAR2	20		
SUB_UNIDADE_ID	SIGLA de subunidade	SUB_UNIDADE	BK	VARCHAR2	20		
SUB_UNIDADE	Designação da subunidade	SUB_UNIDADE		VARCHAR2	200		
TIPO_SUB_UNIDADE	Tipo da subunidade, departamento, serviço, subunidade,....	SUB_UNIDADE		VARCHAR2	20		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.1	<i>Data</i>	05/05/2013	<i>Fontes</i>
<i>Classe</i>	Classificação de um documento segunda a CDU, ISAD, etc...	Tipo1	<i>Hierarquia</i>	N0<N1<N2<N3			
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	ALEPH; Inarte
ID	Chave da dimensão		PK	NUMBER			
CN0	Código nível 0	NÍVEL 0	BK	VARCHAR2	500		
DN0	Designação nível 0	NÍVEL 0		VARCHAR2	500		
CN1	Código nível 1	NÍVEL 1	BK	VARCHAR2	500		
DN1	Designação nível 1	NÍVEL 1		VARCHAR2	500		
TN1	Tipo de nível 1	NÍVEL 1		VARCHAR2	500		
CN2	Código nível 2	NÍVEL 2	BK	VARCHAR2	500		
DN2	Designação nível 2	NÍVEL 2		VARCHAR2	500		
TN2	Tipo de nível 2	NÍVEL 2		VARCHAR2	500		
CN3	Código nível 3	NÍVEL 3	BK	VARCHAR2	500		
DN3	Designação nível 3	NÍVEL 3		VARCHAR2	500		
TN3	Tipo de nível 3	NÍVEL 3		VARCHAR2	500		

Nome	Descrição	SCD	Versão	1.0	Data	05/05/2013	Fontes
Autor	Criador de um documento	Tipo1	Hierarquia				
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	ALEPH; Inarte;Gisa
ID	Chave da dimensão		PK	NUMBER			
Autor	Nome do autor		BK	VARCHAR2	1000		

Nome	Descrição	SCD	Versão	1.0	Data	25/05/2013	Fontes
Categoria Pedido	Categorias de pedidos de apoio de 1ª linha	Tipo 1	Hierarquia	CATEGORIA>SUBCATEGORIA			
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	Sharepoint
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
CATEGORIA	ID da categoria de topo	CATEGORIA	BK	VARCHAR2	50		
DESCRICAO_CATEGORIA	Descrição da categoria de topo	CATEGORIA		VARCHAR2	240		
SUBCATEGORIA	ID do assunto base	SUBCATEGORIA	BK	VARCHAR2	100		
DESCRICAO_SUBCATEGORIA	Descrição do assunto base	SUBCATEGORIA		VARCHAR2	240		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	26/11/2012	<i>Fontes</i>
Site Temático	Sítio web construído propositadamente para um determinado evento	Tipo 1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	<i>Drupal; Seagull; MediaWiki; OJS; Google Analytics; Fontes ad hoc</i>
<i>A ser desenvolvido futuramente...</i>							

Nome	Descrição	SCD	Versão	1.0	Data	27/11/2012	Fontes
Serviço WEB	Qualquer serviço ao utilizador, prestado em linha, gerido ou desenvolvido pela ESE	Tipo 1	Hierarquia				
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	Sharepoint; Redmine; Fontes ad hoc
A ser desenvolvido futuramente...							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
Aplicação de suporte	São todas as aplicações informáticas de suporte à atividade principal do SDI'	Tipo 1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							<i>Aleph; Digitool; Metalib; SFX Inarte; Gisa; Sharepoint; Redmine;</i>

Nome	Descrição	SCD	Versão	1.0	Data	26/05/2013	Fontes
Recurso de pesquisa	Recurso de pesquisa contratado pela FEUP, para apoio às atividades de I&D. Agregadores de informação (BD Bibliográficas) e Editores de conteúdos digitais (<i>Ebooks</i> , Revistas, normas, Patentes)	Tipo 1	Hierarquia	Tipo>Recurso			
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	Metalib; SFX; Recursos eletrónicos; Fontes ad hoc
DIMENSION_KEY	Chave da dimensão		PK	NUMBER			
TIPO_RECURSO_ID	ID do nível de topo	TIPO	BK	VARCHAR2	4		
DESC_TIPO_RECURSO	Designação da tipologia do recurso	TIPO		VARCHAR2	50		
RECURSO_ID_1	ID do recurso	RECURSO	BK	VARCHAR2	8		
NOME_RECURSO	Nome do recurso	RECURSO		VARCHAR2	100		
NOME_ABREVIADO	Nome abreviado do recurso	RECURSO		VARCHAR2	100		
TIPO_SUBSCRICAO	Tipo de subscrição [Gratuito Pago]	RECURSO		VARCHAR2	20		
EDITOR	Editor do recurso	RECURSO		VARCHAR2	100		
CATEGORIAS	Texto livro com alguma caracterização do recurso	RECURSO		VARCHAR2	50		
NOMES_ADICIONAIS	Identificadores adicionais para o recurso	RECURSO		VARCHAR2	100		
HOME_PAGE	Sítio web base do recurso	RECURSO		VARCHAR2	100		
CRIADOR	Entidade autora do recurso	RECURSO		VARCHAR2	100		
TIPO	Outro campo com alguma caracterização, não estruturada do recurso	RECURSO		VARCHAR2	50		
URL	Ligação para funcionalidade de pesquisa	RECURSO		VARCHAR2	100		
HOSTNAME_PORTA	Parâmetros para a configuração do recurso (pode ter interesse)	RECURSO		VARCHAR2	100		
CODIGO_BD	Códigos internos do recurso (pode ter interesse)	RECURSO		VARCHAR2	50		

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
Funcionalidade	É uma funcionalidade gerida ou desenvolvida pela ESE, sempre dentro do contexto das aplicações de suporte ou serviços WEB disponibilizados pelo SDI e usados pelos utilizadores. (ex. sistema de rating, sistema de renovação por sms, etc)	Tipo 1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							<i>Aleph; Digitoool; Metalib; SFX Inarte; Gisa; Sharepoint; Redmine; Fontes ad hoc; Drupal; Seagull; MediaWiki; OJS; Sítio da FEUP Edições</i>

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
Material de comunicação	Todo o material digital ou impresso desenvolvido pela ESE para suporte e apoio das mais diversas iniciativas do SDI (ex. banners, panfletos, sinaléticas cartazes, etc)	Tipo 1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							<i>Fontes ad hoc; Drupal; Seagull; MediaWiki; OJS; Sítio da FEUP Edições</i>

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
Bibliografia UC	Bibliografia principal e secundária das Unidades Curriculares lecionadas na FEUP	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							SIGARRA

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	27/11/2012	<i>Fontes</i>
Unidade Curricular	Disciplina lecionada em pelo menos um curso da FEUP	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							SIGARRA

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Sala	Identificador de uma sala localizada num qualquer piso de um edifício da FEUP	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							SIGARRA; Fontes adhoc

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Estante depósito	Identificador das estantes dos depósitos do SDI	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							GISA; Fontes ad hoc

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Equipamento	Material mobiliário, como por exemplo computadores, cadeiras, secretárias, fotocopiadoras, etc...	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							SIGARRA; ALEPH; Fontes <i>ad hoc</i>

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Incidente	Acontecimento anômalo que provoca alterações ao estado de conservação de espaços e equipamentos do SDI	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							SIGARRA; <i>Sharepoint</i> ; Fontes <i>ad hoc</i>

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Projeto cooperação	Projetos de cooperação e integração nos quais o SDI está envolvido, a nível nacional e a nível internacional	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							Sharepoint; SIGARRA(?)

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Projeto	Planificação de trabalhos, habitualmente geridos com recurso a gestores de projetos e/ou manualmente com diagramas de <i>gantt</i>	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
<i>A ser desenvolvido futuramente...</i>							<i>Sharepoint; Redmine; Fontes ad hoc</i>

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Evento	Acontecimento de cariz cultural ou científico organizado pelo SDI	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	Sharepoint; Fontes <i>ad hoc</i> ; SIGARRA; Eventos
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Instrução de trabalho	Conjunto de procedimentos para executar uma dada tarefa	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	
A ser desenvolvido futuramente...							Sharepoint

Nome	Descrição	SCD	Versão	1.0	Data	28/11/2012	Fontes
Tarefa	Trabalho que se deve fazer dentro de um determinado limite de tempo, pertencente a um projeto mais abrangente	Tipo1	Hierarquia				
Atributo	Descrição	Nível	Chave	Tipo	Tam	Prec	Redmine; Sharepoint; Fontes ad hoc
A ser desenvolvido futuramente...							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Plano de atividade	Plano de atividades anual do SDI	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; Sharepoint
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Orçamento	Verba previamente reservada para a aquisição de bens e serviços	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; ALEPH
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Rubrica	Designação de uma classificação orçamental para despesas ou proveitos de um mesmo tipo	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; Fontes <i>ad hoc</i>
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
CCO	Centro de Controlo Orçamental	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Produto	Bens adquiridos ou vendidos	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; Sharepoint
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
PAD	Pedido de autorização de despesa	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; Sharepoint
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Fornecedor	Entidade responsável pela venda de bens ao SDI	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	SIGARRA; ALEPH; Fontes <i>ad hoc</i>
<i>A ser desenvolvido futuramente...</i>							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Humidade	Quantidade de vapor de água registada em determinado local	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	Sharepoint; Fontes ad hoc
A ser desenvolvido futuramente...							

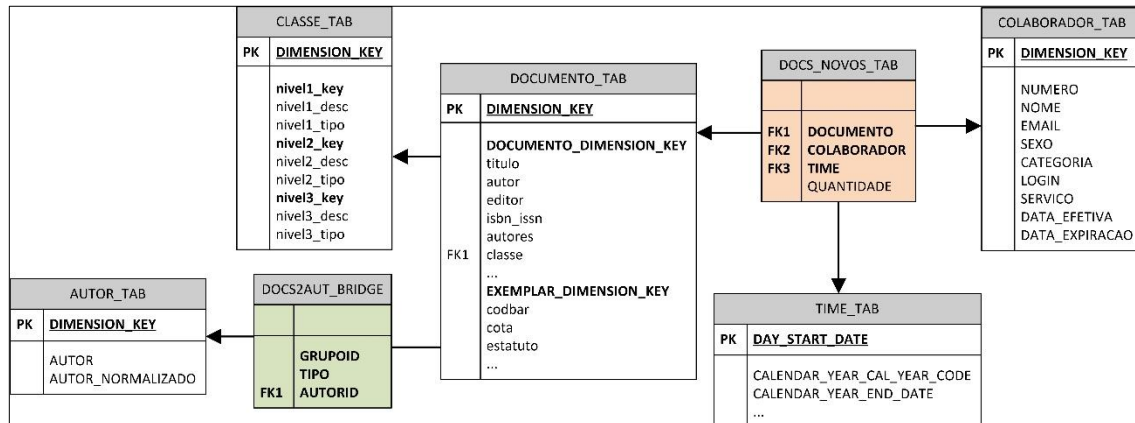
<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Temperatura	Grau de aquecimento ou arrefecimento de um local	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	Sharepoint; Fontes ad hoc
A ser desenvolvido futuramente...							

<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
Luminosidade	Quantidade de luz medida em determinado local	Tipo1	<i>Hierarquia</i>				
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	Sharepoint; Fontes ad hoc
A ser desenvolvido futuramente...							

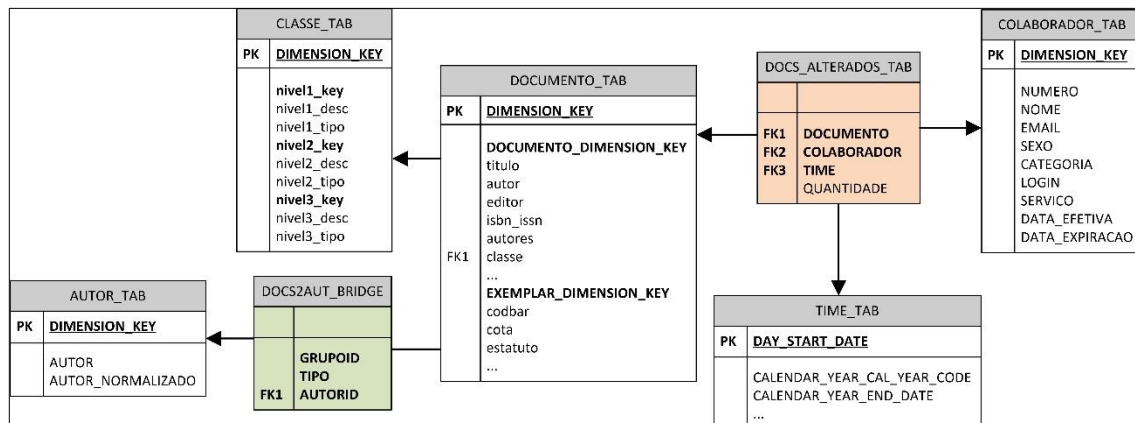
<i>Nome</i>	<i>Descrição</i>	<i>SCD</i>	<i>Versão</i>	1.0	<i>Data</i>	28/11/2012	<i>Fontes</i>
<i>Tempo</i>	Momento no qual determinado evento ocorre	Tipo1	<i>Hierarquia</i>	Ano<Mês<Dia			
<i>Atributo</i>	<i>Descrição</i>	<i>Nível</i>	<i>Chave</i>	<i>Tipo</i>	<i>Tam</i>	<i>Prec</i>	Sharepoint; Fontes <i>ad hoc</i> ; ALEPH;Digitool;Metalib; SFX; Recursos eletrónicos;...
Implementado o modelo pré-determinado do <i>Oracle Warehouse Builder</i>							

Anexo H. Diagramas Entidade-Relação

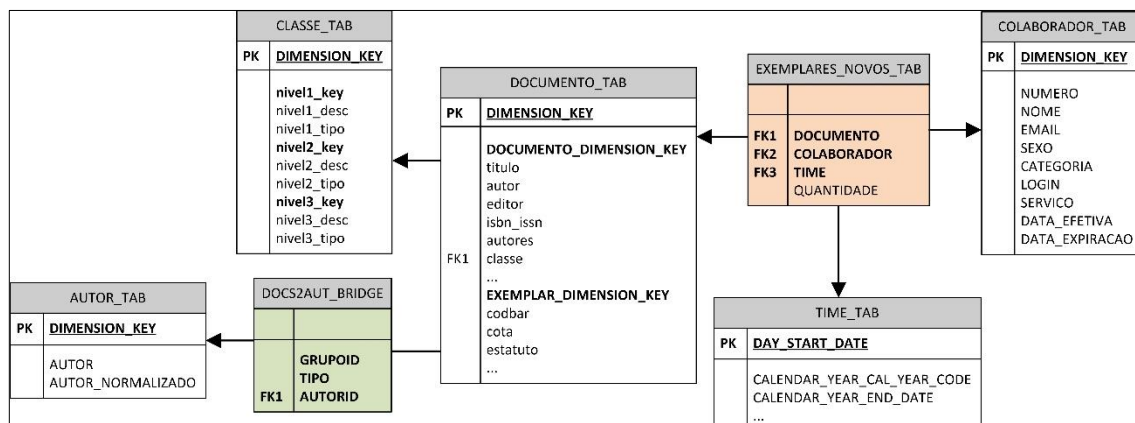
H.1 Estrela - Documentos novos



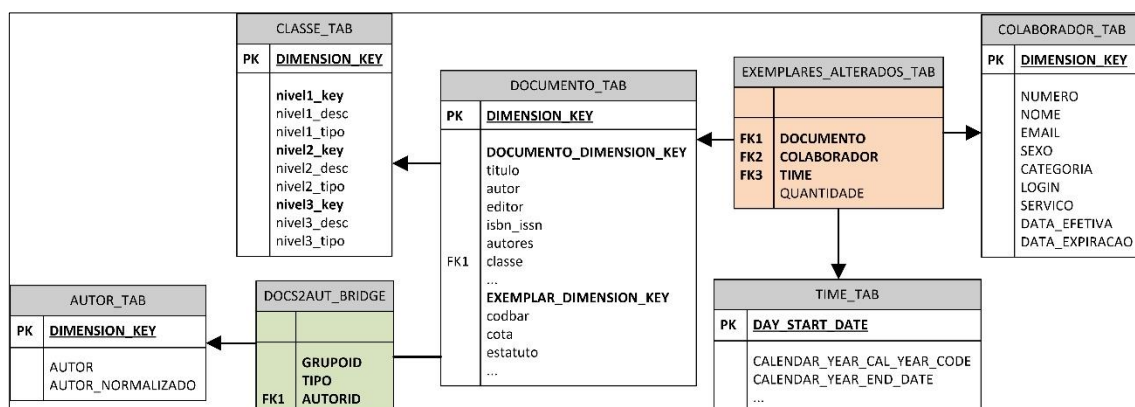
H.2 Estrela - Documentos Alterados



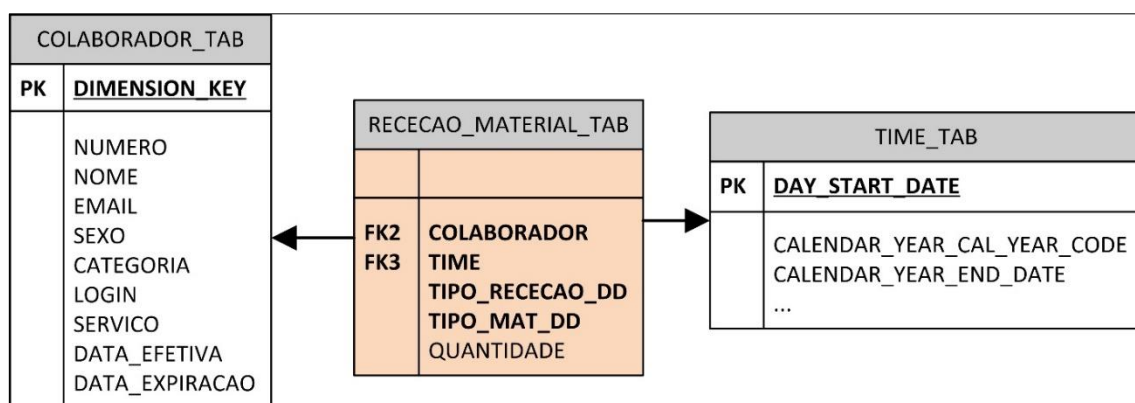
H.3 Estrela – Exemplos alterados



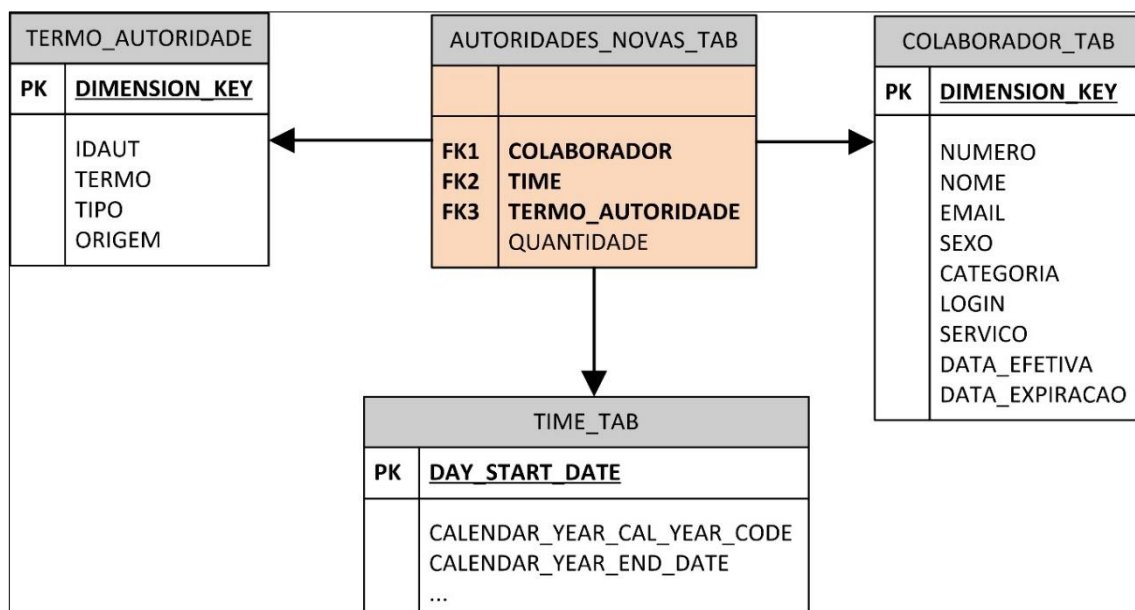
H.4 Estrela – Exemplos alterados



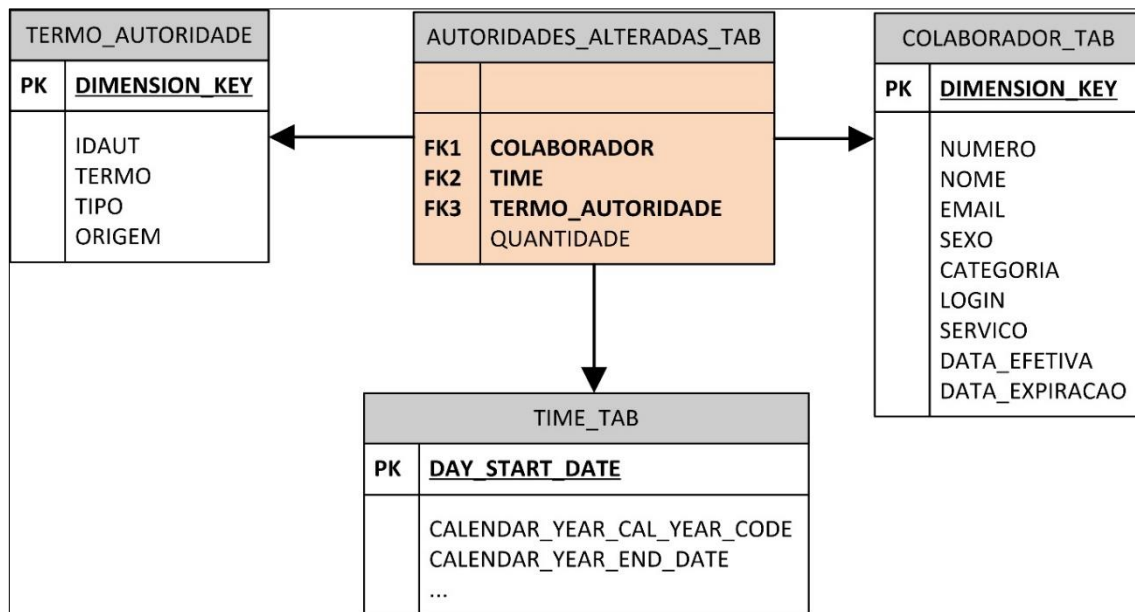
H.5 Estrela - Receção de material



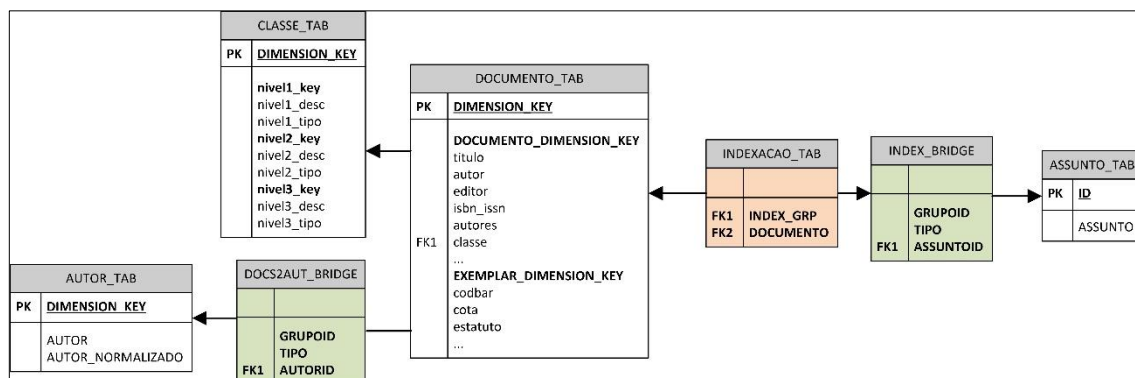
H.6 Estrela - Autoridades novas



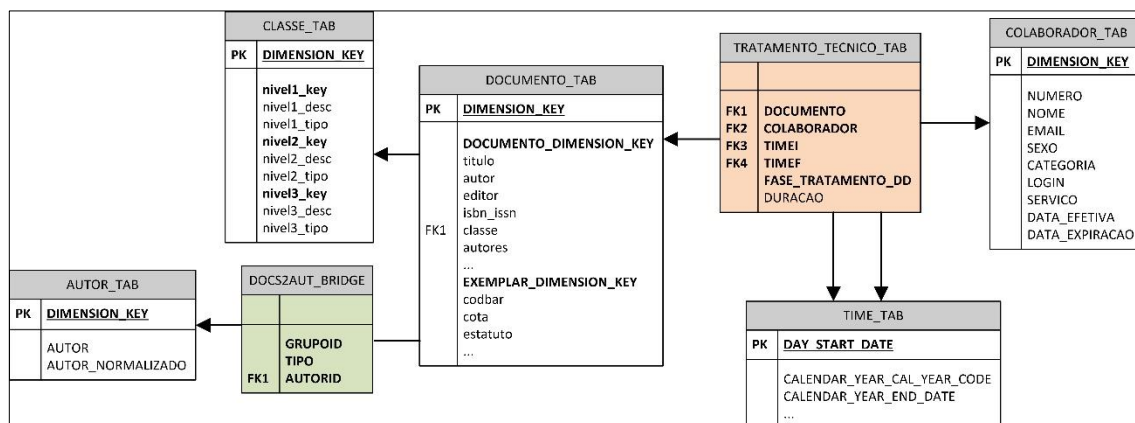
H.7 Estrela – Autoridades alteradas



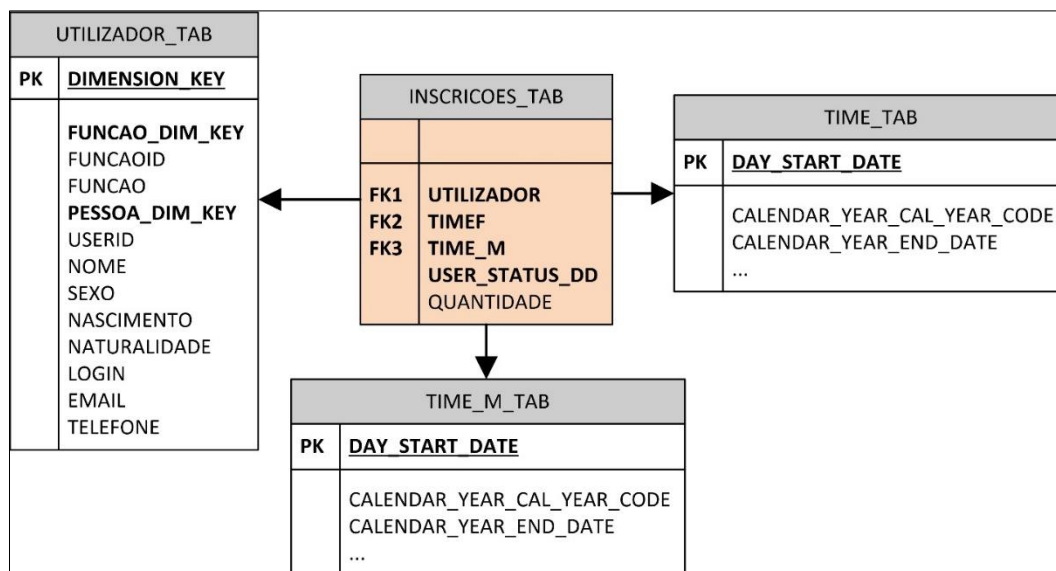
H.8 Estrela - Indexação



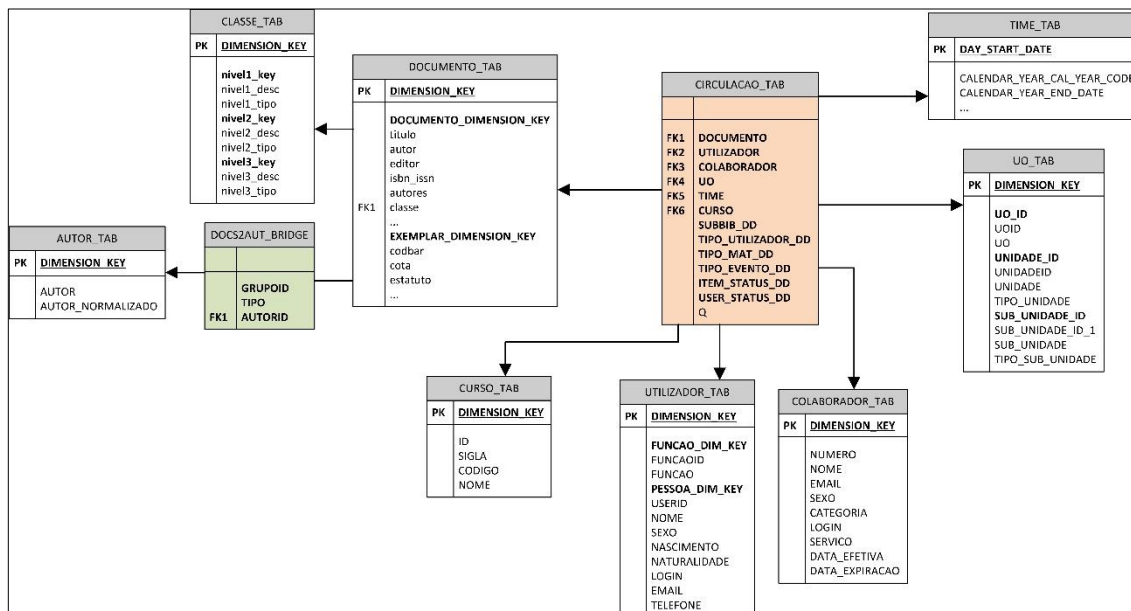
H.9 Estrela - Tratamento técnico



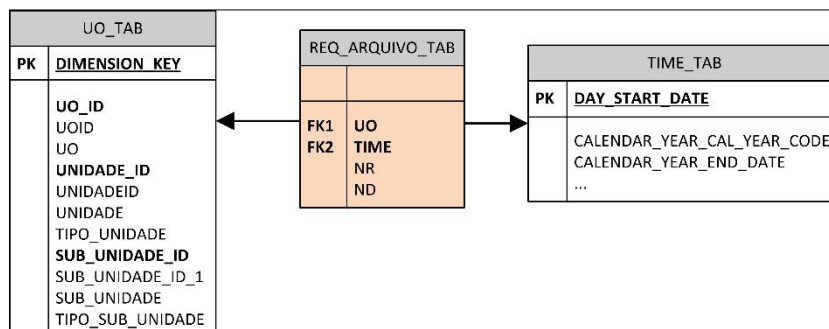
H.10 Estrela - Inscrições



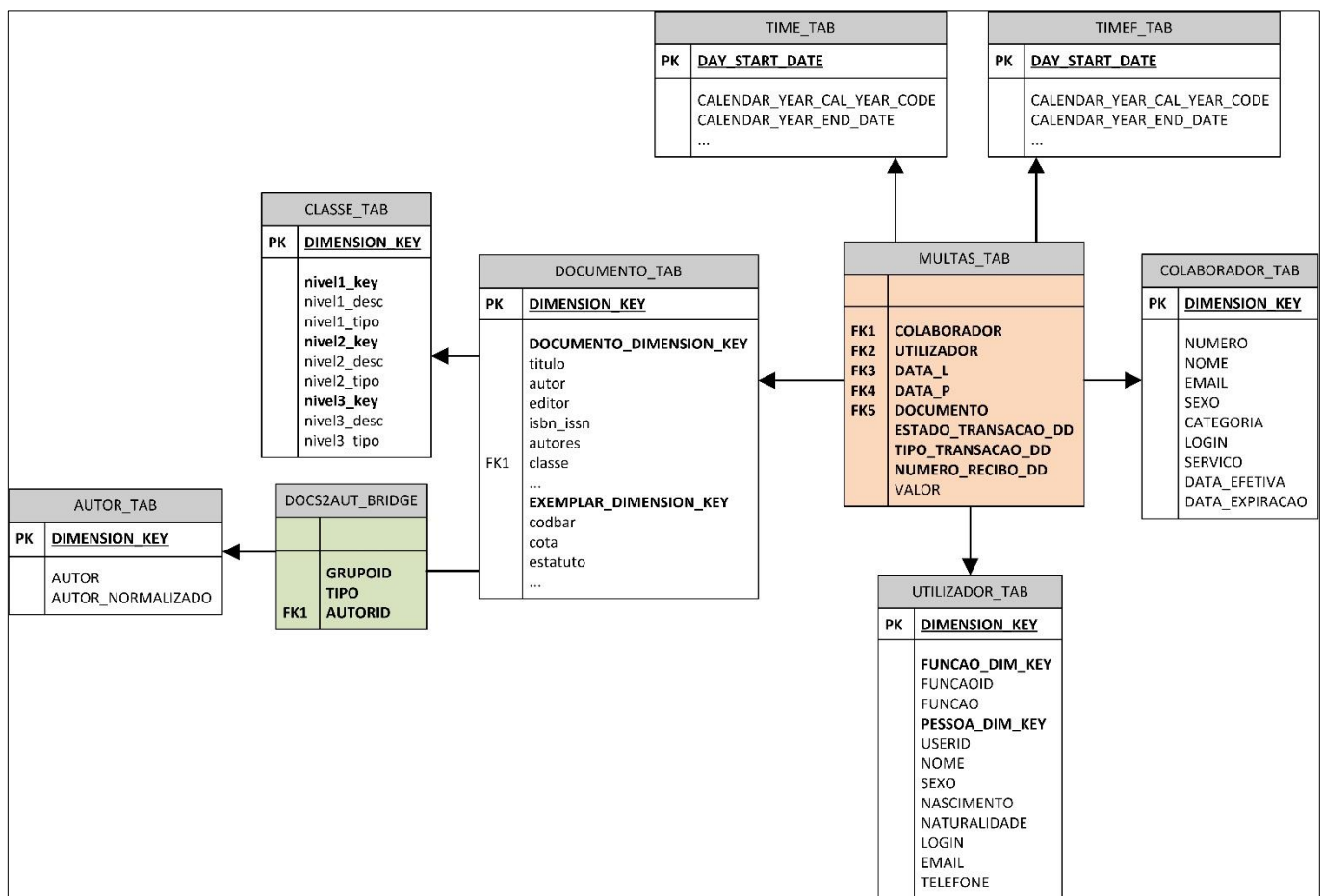
H.11 Estrela - Circulação



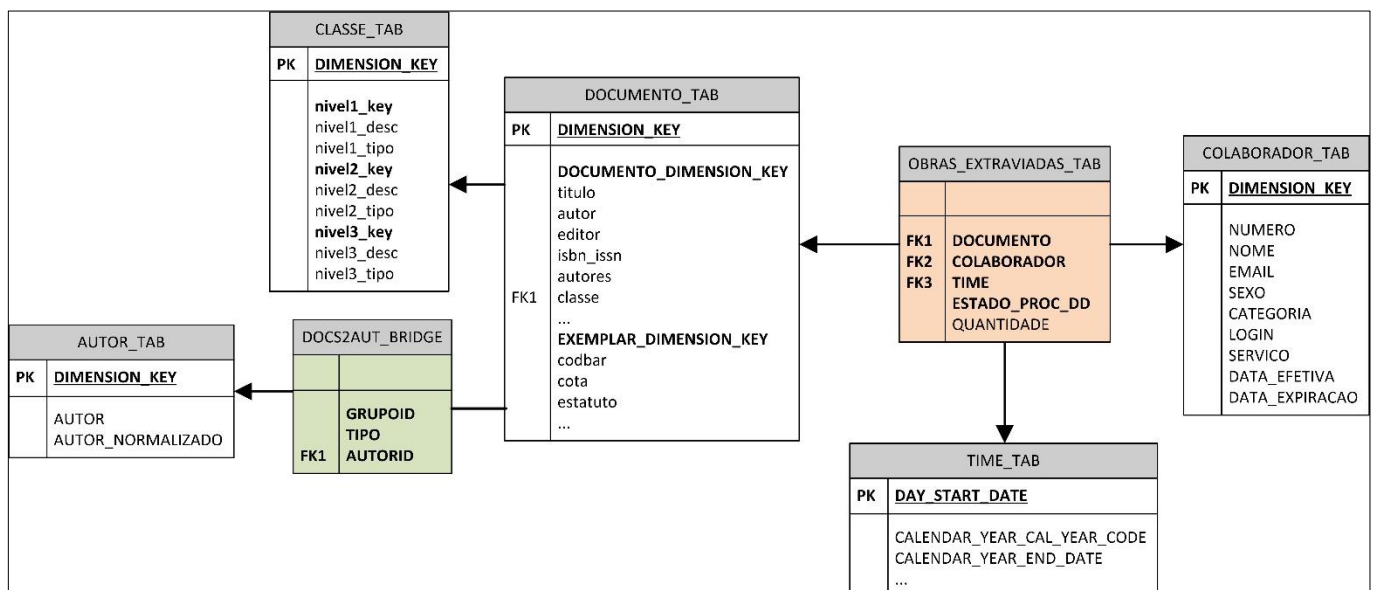
H.12 Estrela – Requisições ao arquivo



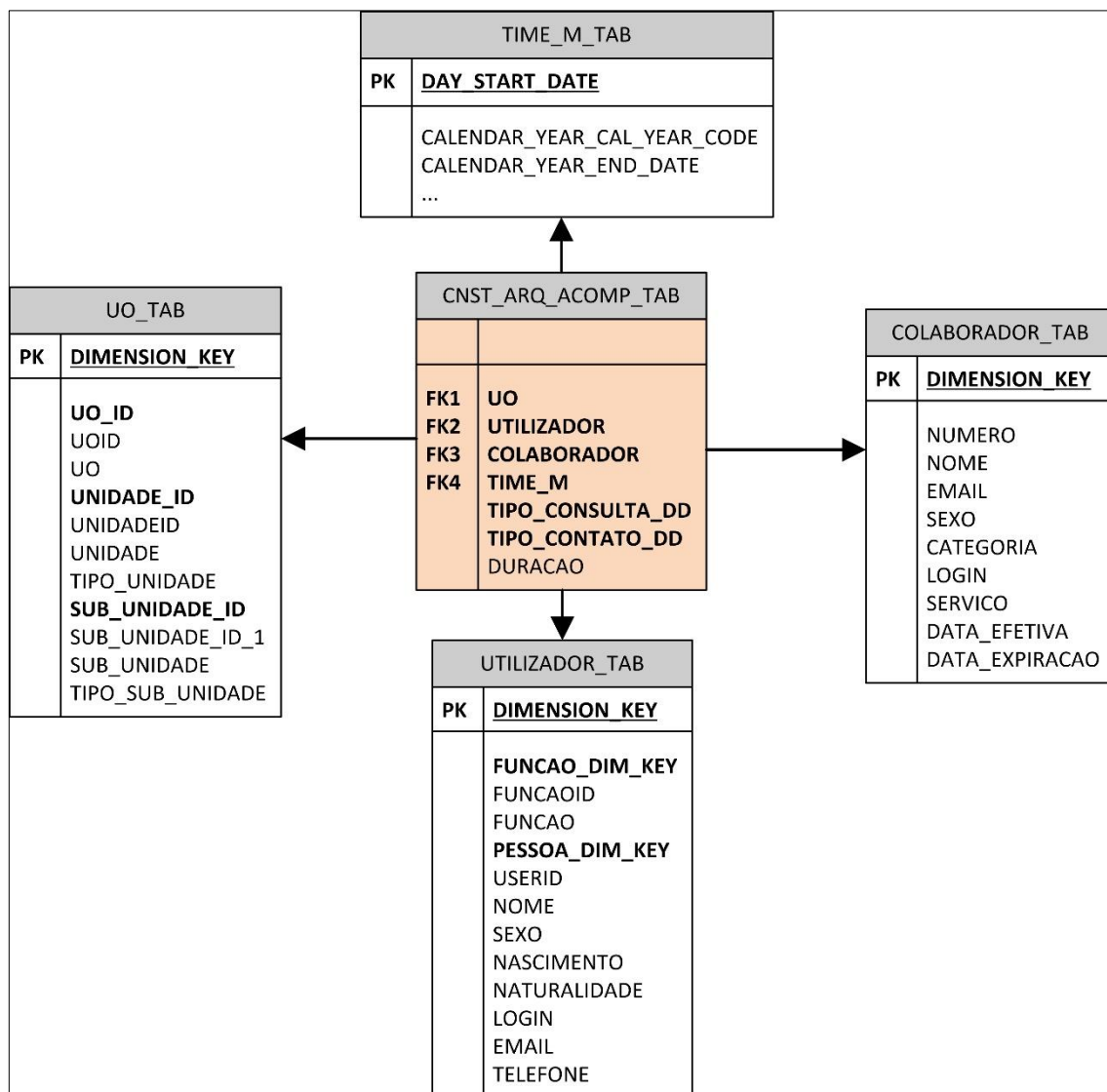
H.13 Estrela - Multas



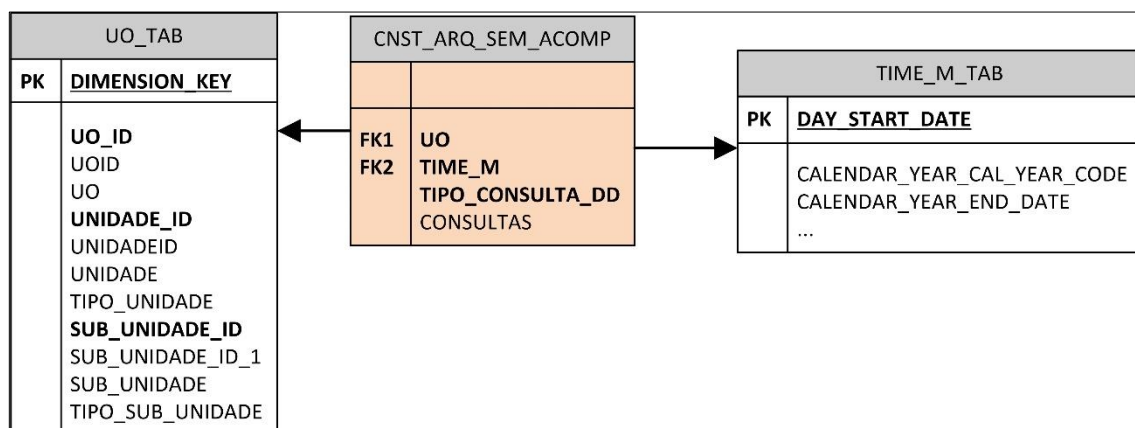
H.14 Estrela – Obras extraviadas



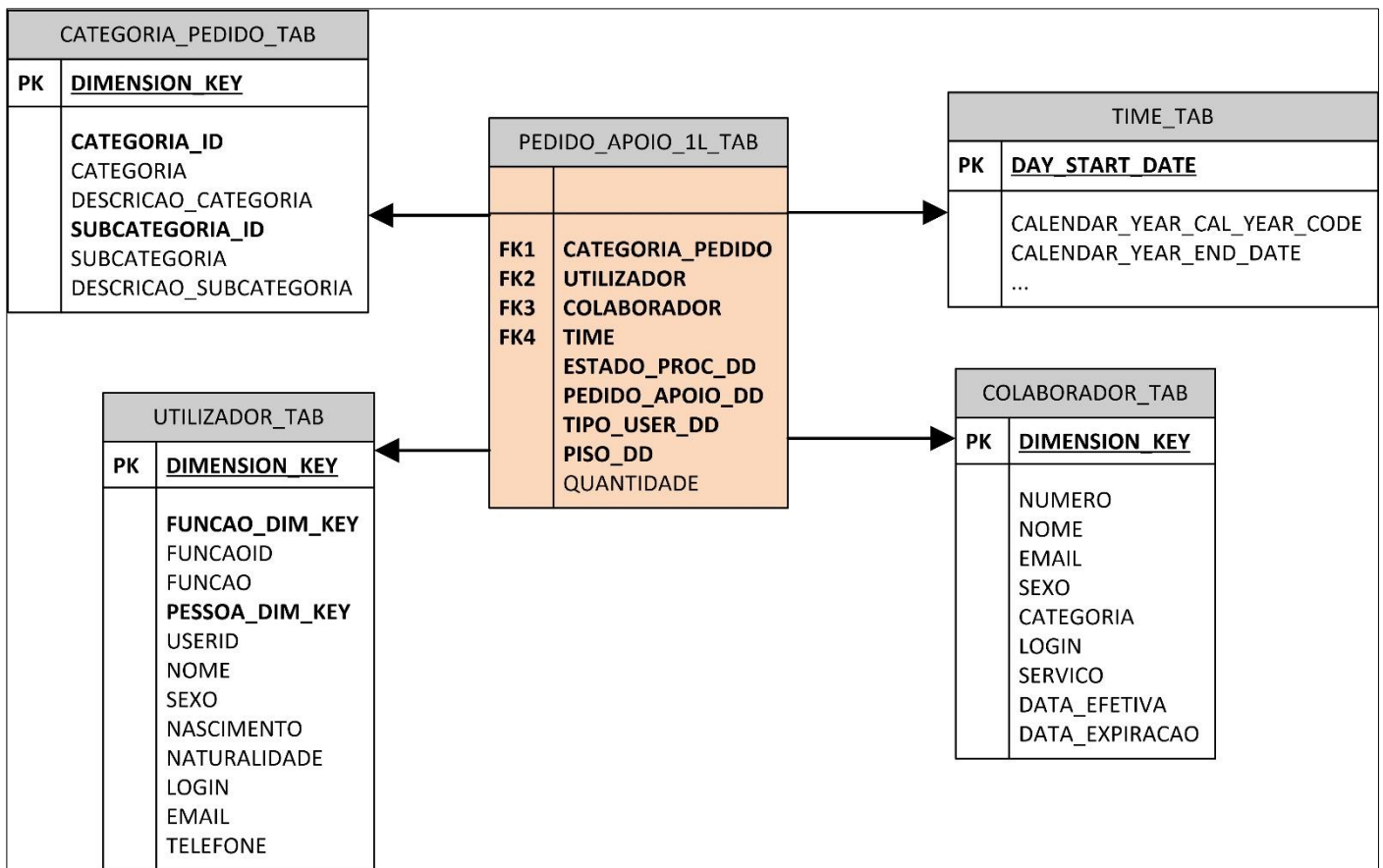
H.15 Estrela – Consultas ao arquivo acompanhadas



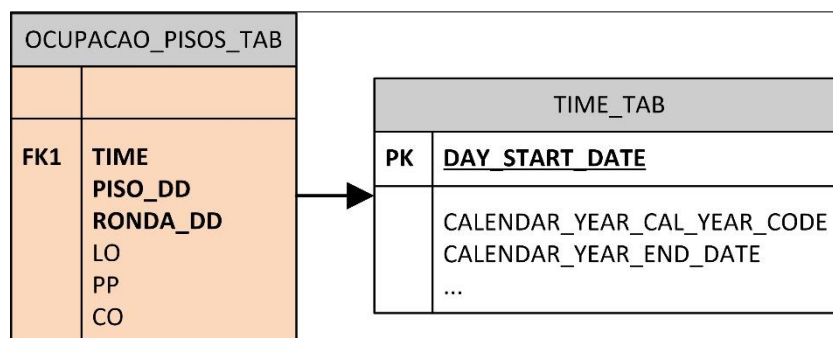
H.16 Estrela – Consultas ao arquivo sem acompanhamento



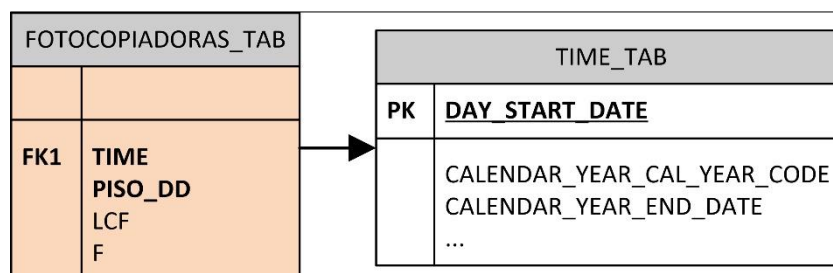
H.17 Estrela – Pedidos de apoio de 1.ª linha



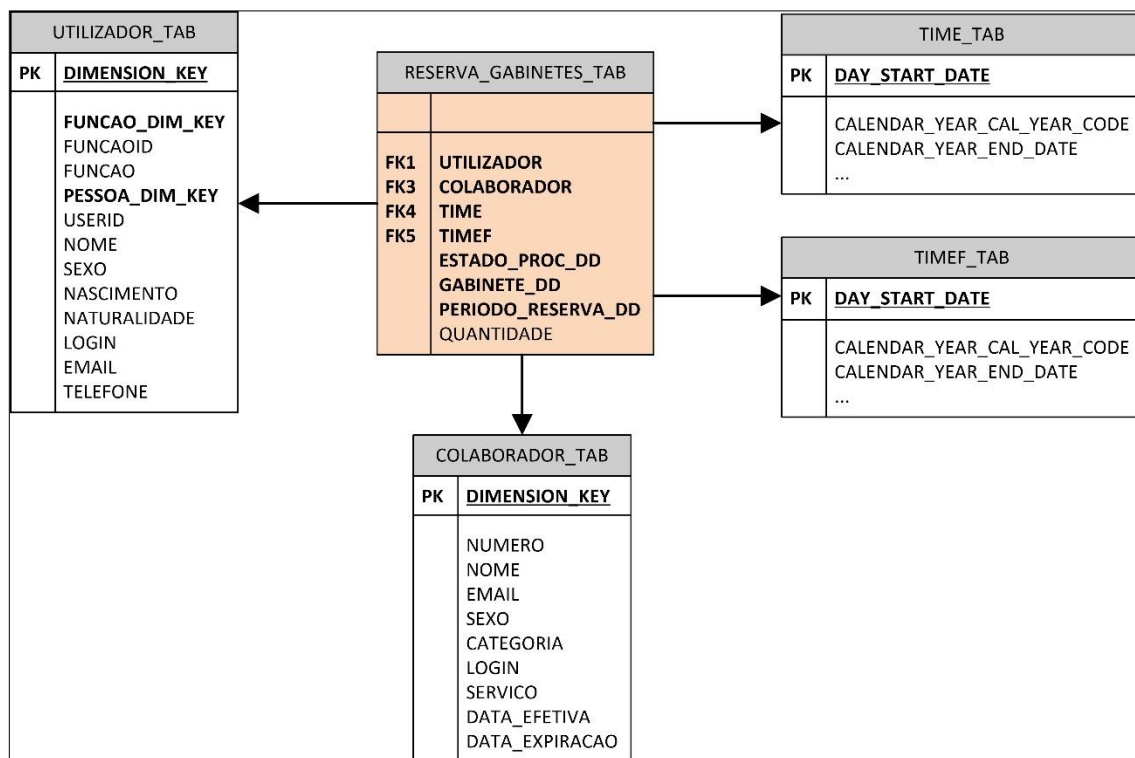
H.18 Estrela – Ocupação dos pisos



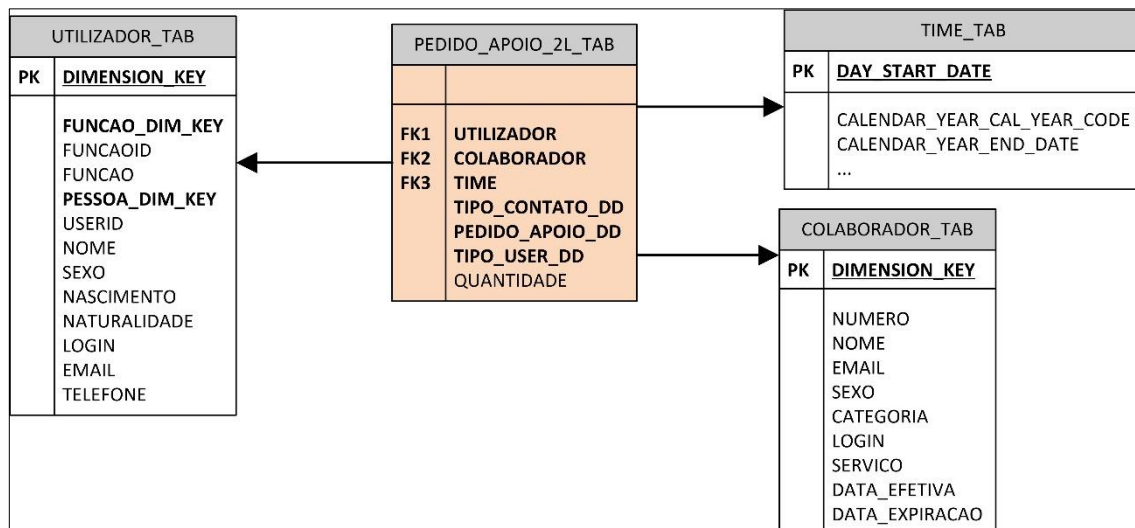
H.19 Estrela - Fotocopiadoras



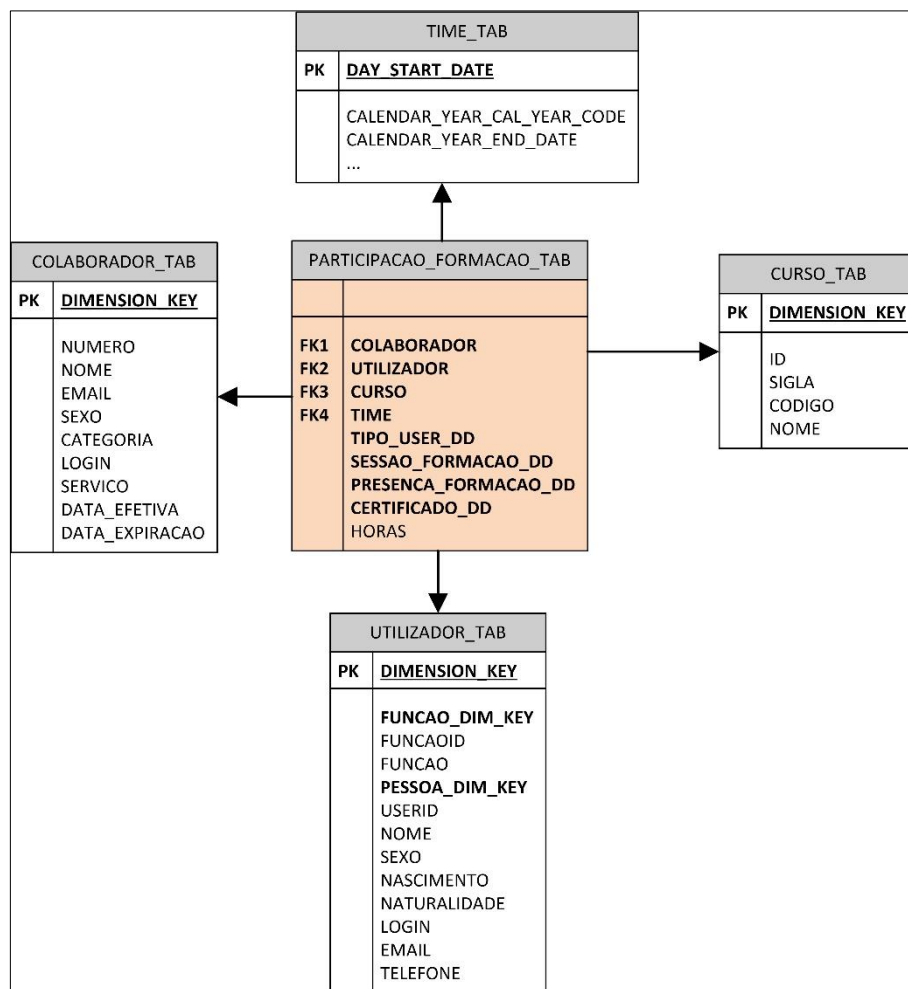
H.20 Estrela – Reserva de gabinetes



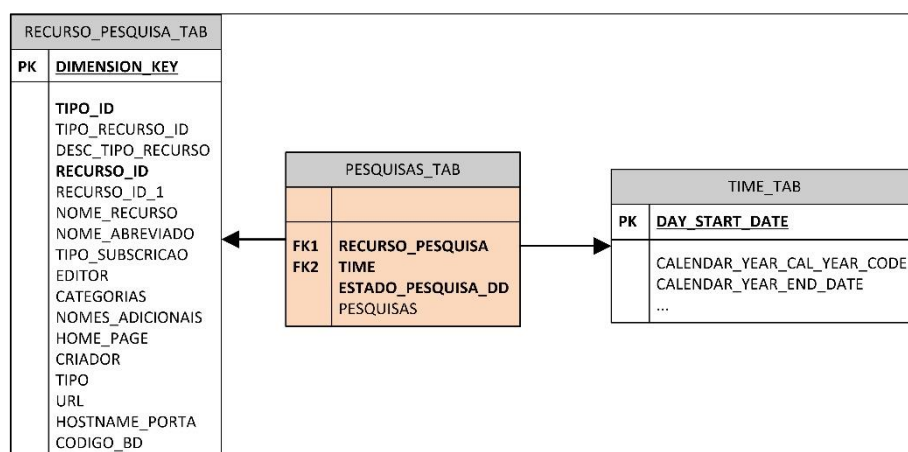
H.21 Estrela – Pedidos de apoio de 2.ª linha



H.22 Estrela – Participação em formação



H.23 Estrela - Pesquisas



Anexo I. Código PL/SQL

I.1 Código para preencher uma tabela de ponte

```
CREATE OR REPLACE
PACKAGE BODY "FILL_DOC2AUT_BRIDGE" AS

PROCEDURE "FILL_DOC2AUTHORS_BRIDGE"("IN_DOCID" IN VARCHAR2, "IN_OVERRIDE" IN VARCHAR2) IS
--initialize variables here
CURSOR AUTORES_DOCS_CURSOR IS
  SELECT DOCID, TIPO, AUTOR FROM EXT_AUTORES_DOCS WHERE DOCID like IN_DOCID;
REC_CURRENT_DOC  AUTORES_DOCS_CURSOR%ROWTYPE;
CHECK_BRIDGE     NUMBER;
GROUP_KEY        EXT_AUTORES_GRUPOS.ID_GRUPO%TYPE;
AUTOR_KEY        AUTOR_TAB.ID%TYPE;
DOCID_KEY        DOCUMENTO_TAB.DOC_ID%TYPE;
AUTOR_STATUS     DOCUMENTO_TAB.AUTOR%TYPE;
-- main body
BEGIN
  DBMS_OUTPUT.ENABLE(NULL);
  OPEN AUTORES_DOCS_CURSOR;
  LOOP
    FETCH AUTORES_DOCS_CURSOR INTO REC_CURRENT_DOC;
    EXIT WHEN AUTORES_DOCS_CURSOR%NOTFOUND;
    BEGIN
      SELECT DISTINCT AUTOR INTO AUTOR_STATUS
      FROM DOCUMENTO_TAB
      WHERE DOC_ID = REC_CURRENT_DOC.DOCID;
      IF (AUTOR_STATUS IS NOT NULL AND UPPER(IN_OVERRIDE) = 'N') THEN
        DBMS_OUTPUT.PUT_LINE('[[DEBUG]Documento '||REC_CURRENT_DOC.DOCID||' protegido contra escrita');
      ELSE
        SELECT ID_GRUPO INTO GROUP_KEY
        FROM EXT_AUTORES_GRUPOS, EXT_DOCS_STRAGG
        WHERE EXT_DOCS_STRAGG.DOCID = REC_CURRENT_DOC.DOCID
        AND EXT_DOCS_STRAGG.GRUPO = EXT_AUTORES_GRUPOS.GRUPO;
        IF GROUP_KEY IS NOT NULL THEN
          DBMS_OUTPUT.PUT_LINE('[[DEBUG]A processar documento '||REC_CURRENT_DOC.DOCID||' : Grupo ->
'||GROUP_KEY);
          SELECT AUTOR_TAB.ID INTO AUTOR_KEY
          FROM AUTOR_TAB
          WHERE AUTOR_TAB.AUTOR = REC_CURRENT_DOC.AUTOR;
          IF AUTOR_KEY IS NOT NULL THEN
            DBMS_OUTPUT.PUT_LINE('[[DEBUG]A processar documento '||REC_CURRENT_DOC.DOCID||' : AUTOR ->
'||AUTOR_KEY);
            SELECT DISTINCT DOC_ID INTO DOCID_KEY
            FROM DOCUMENTO_TAB
            WHERE DOC_ID = REC_CURRENT_DOC.DOCID;
            IF DOCID_KEY IS NOT NULL THEN
              DBMS_OUTPUT.PUT_LINE('[[DEBUG]A atualizar dimensão Documento: '||REC_CURRENT_DOC.DOCID);
              UPDATE DOCUMENTO_TAB
              SET AUTOR = GROUP_KEY
              WHERE DOCUMENTO_TAB.DOC_ID = REC_CURRENT_DOC.DOCID;
              SELECT COUNT(*) INTO CHECK_BRIDGE
              FROM DOC2AUT_BRIDGE
              WHERE GRUPOID = GROUP_KEY
              AND TIPO= REC_CURRENT_DOC.TIPO
              AND AUTORID=AUTOR_KEY;
              -- DBMS_OUTPUT.PUT_LINE('[[DEBUG] CHECK_BRIDGE: '||CHECK_BRIDGE);
            
```

```

    IF CHECK_BRIDGE = 0 THEN
        DBMS_OUTPUT.PUT_LINE('[[DEBUG] A inserir '||GROUP_KEY||' na tabela de ponte...');
        INSERT INTO DOC2AUT_BRIDGE (GRUPOID, TIPO, AUTORID)
        VALUES ( GROUP_KEY, REC_CURRENT_DOC.TIPO, AUTOR_KEY);
    ELSE
        DBMS_OUTPUT.PUT_LINE('[[DEBUG] Conjunto <Grupo,Tipo,Autor> já existem na tabela de ponte. Tabela de Ponte
não será atualizada. ');
    END IF;
    COMMIT;
END IF;
END IF;
END IF;
END IF;
EXCEPTION
    WHEN NO_DATA_FOUND
    THEN
        DBMS_OUTPUT.PUT_LINE('[[DEBUG] Autores do documento '||REC_CURRENT_DOC.DOCID||' não foram actualizados');
    END;
END LOOP;
CLOSE AUTORES_DOCS_CURSOR;
EXCEPTION
    WHEN OTHERS THEN
        DBMS_OUTPUT.PUT_LINE('Operation failed '||SQLCODE: '||SQLCODE);
        DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
        -- enter any exception code here
END;
-- End of FILL_DOC2AUTHORS_BRIDGE;
END "FILL_DOC2AUT_BRIDGE";

```

I.2 Código alternativo para preenchimento de uma tabela de factos

```
CREATE OR REPLACE
PACKAGE BODY "INSERT_FACTS" AS

PROCEDURE "DOC_ALT_FACTS" IS
--initialize variables here
CURSOR doc_alterado_facts_cursor IS
  select docid, catalogador, data from ext_factos_docs_alt_bib;-- where substr(to_char(data,'dd-mm-yyyy'),7,4) = 2012;

rec_cur_fact  doc_alterado_facts_cursor%ROWTYPE;

doc_dim_key documento_tab.DOCUMENTO_DIMENSION_KEY%type;
colab_dim_key colaborador_tab.dim_key%type;
time_dim_key  time_tab.day_start_date%type;

-- main body
BEGIN
  DBMS_OUTPUT.ENABLE(NULL);
  OPEN doc_alterado_facts_cursor;
  LOOP
    FETCH doc_alterado_facts_cursor INTO rec_cur_fact;
    EXIT WHEN doc_alterado_facts_cursor%NOTFOUND;

    BEGIN
      SELECT documento_tab.DOCUMENTO_DIMENSION_KEY INTO doc_dim_key
      FROM documento_tab
      WHERE rec_cur_fact.docid = documento_tab.doc_id AND
      documento_tab.DOCUMENTO_DIMENSION_KEY = documento_tab.DIMENSION_KEY;
      IF doc_dim_key IS NOT NULL THEN
        SELECT colaborador_tab.dim_key INTO colab_dim_key
        FROM colaborador_tab
        WHERE rec_cur_fact.catalogador=colaborador_tab.login;
        IF colab_dim_key IS NOT NULL THEN
          SELECT time_tab.day_start_date INTO time_dim_key
          FROM time_tab
          WHERE rec_cur_fact.data = time_tab.day_start_date;
          IF time_dim_key IS NOT NULL THEN
            INSERT INTO DOCS_ALTERADOS_TAB (QUANTIDADE, COLABORADOR, DOCUMENTO, TIME)
            VALUES ( 1, colab_dim_key, doc_dim_key, time_dim_key);
          END IF;
        END IF;
      END IF;
    EXCEPTION
      WHEN NO_DATA_FOUND
      THEN
        DBMS_OUTPUT.PUT_LINE('[[DEBUG]]Documento '||rec_cur_fact.docid||' não encontrado na dimensão');
      END;

  END LOOP;
```

```
CLOSE doc_alterado_facts_cursor;  
EXCEPTION  
  WHEN OTHERS THEN  
    DBMS_OUTPUT.PUT_LINE('Operation failed '||SQLCODE: '||SQLCODE);  
    DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);  
    -- enter any exception code here  
END;  
-- End of DOC_ALT_FACTS;  
END "INSERT_FACTS";
```

I.3 Código para retificar números de alunos

```
CREATE OR REPLACE
PACKAGE BODY NALU AS

PROCEDURE CORRIGE_Z35 AS

    CURSOR OLDSTUDENTNUMBERSCURSOR IS
    SELECT * FROM NEWNALUNOS WHERE VELHOID IN (SELECT DISTINCT TRIM(USERID) FROM
EXT_EUP50Z35_2011 WHERE USERID LIKE '100578006%');

    CURRENTOLDID NEWNALUNOS%ROWTYPE;

BEGIN

    DBMS_OUTPUT.ENABLE(NULL);

    OPEN OLDSTUDENTNUMBERSCURSOR;

    LOOP
        FETCH OLDSTUDENTNUMBERSCURSOR INTO CURRENTOLDID;
        EXIT WHEN OLDSTUDENTNUMBERSCURSOR%NOTFOUND;

        UPDATE EXT_EUP50Z35_2011 SET USERID = CURRENTOLDID.NOVOID WHERE USERID LIKE
CURRENTOLDID.VELHOID||'%';

        DBMS_OUTPUT.PUT_LINE('[[DEBUG]]A ATUALIZAR ID ANTIGO: '||CURRENTOLDID.VELHOID);

        COMMIT;
    END LOOP;
    CLOSE OLDSTUDENTNUMBERSCURSOR;
    EXCEPTION
    WHEN OTHERS THEN
        DBMS_OUTPUT.PUT_LINE('Operation failed '||SQLCODE: '||SQLCODE);
        DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
    END CORRIGE_Z35;

PROCEDURE CORRIGE_INSCRICOES AS

    CURSOR OLDSTUDENTNUMBERSCURSOR IS
    SELECT * FROM NEWNALUNOS WHERE VELHOID IN (SELECT DISTINCT USERID FROM EXT_INSCRICOES WHERE
USERID LIKE '%');

    CURRENTOLDID NEWNALUNOS%ROWTYPE;

BEGIN

    DBMS_OUTPUT.ENABLE(NULL);

    OPEN OLDSTUDENTNUMBERSCURSOR;
```

```

OPEN OLDSTUDENTNUMBERSCURSOR;

LOOP
  FETCH OLDSTUDENTNUMBERSCURSOR INTO CURRENTOLDID;
  EXIT WHEN OLDSTUDENTNUMBERSCURSOR%NOTFOUND;

  UPDATE EXT_INSCRICOES SET USERID = CURRENTOLDID.NOVOID WHERE USERID LIKE
CURRENTOLDID.VELHOLD||'%';

  DBMS_OUTPUT.PUT_LINE('[[DEBUG]A ATUALIZAR ID ANTIGO: '||CURRENTOLDID.VELHOLD);

  COMMIT;
END LOOP;

CLOSE OLDSTUDENTNUMBERSCURSOR;

EXCEPTION
WHEN OTHERS THEN
  DBMS_OUTPUT.PUT_LINE('Operation failed '||SQLCODE: '||SQLCODE);
  DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);

END CORRIGE_INSCRICOES;
END NALU;

```


I.4 Código para normalizar visualização de autores

```
CREATE OR REPLACE
PACKAGE BODY "NORMALIZA_AUTORES" AS

PROCEDURE "NORMALIZA_AUTORES"("V_NOME" IN VARCHAR2 DEFAULT '%') IS
--initialize variables here
CURSOR AUTORES_CURSOR IS
  SELECT * FROM AUTOR_TAB WHERE AUTOR_NORMALIZADO like v_nome;

REC_CURRENT_AUTOR  AUTOR_TAB%ROWTYPE;

CLEANED_NAME      AUTOR_TAB.AUTOR_NORMALIZADO%TYPE;

-- main body
BEGIN
  DBMS_OUTPUT.ENABLE(NULL);
  OPEN AUTORES_CURSOR;
  LOOP
    FETCH AUTORES_CURSOR INTO REC_CURRENT_AUTOR;
    EXIT WHEN AUTORES_CURSOR%NOTFOUND;
    BEGIN
      DBMS_OUTPUT.PUT_LINE('[[DEBUG]A processar Autor '||REC_CURRENT_AUTOR.ID);
      -- início do processo de normalização dos nomes dos autores
      CLEANED_NAME := REC_CURRENT_AUTOR.AUTOR_NORMALIZADO;
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=and>>', 'and');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<Mc=Mac>>', 'Mac');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<è=e>>', 'è');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<(12=12)>>', '12');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<(28=28)>>', '28');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<4=4>>', '4');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<12=12>>', '12');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<28=28>>', '28');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<82=82>>', '82');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<83=83>>', '83');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<90=90>>', '90');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<219=219>>', '219');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<319=319>>', '319');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<348=348>>', '348');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<318=318>>', '318');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<1920=1920>>', '1920');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<(1920=1920)>>', '1920');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<1947=1947>>', '1947');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<1962=1962>>', '1962');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<1988=1988>>', '1988');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<2010=2010>>', '2010');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<2000=2000>>', '2000');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<2001=2001>>', '2001');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<2003=2003>>', '2003');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<2004=2004>>', '2004');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<1963-1964=1963- 1964>>', '1963-1964');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=e>>', '&');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=und>>', '&');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<è=>>', 'è');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=>>', 'è');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<I=I>>', 'I');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<II=2>>', 'II');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<III=3>>', 'III');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<VI=6>>', 'VI');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<X=10>>', 'X');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<XI=11>>', 'XI');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<XXI=21>>', 'XXI');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=et>>', '&');
      CLEANED_NAME := REPLACE(CLEANED_NAME, '<<=&=andZ>>', '&');
```

```

CLEANED_NAME := REPLACE(CLEANED_NAME,'<<1971=One Thousand, Nine Hundred Seventy One>>','1971');
CLEANED_NAME := REPLACE(CLEANED_NAME,'<<E10=E10>>','E10');
CLEANED_NAME := REPLACE(CLEANED_NAME,'<<C 2=C 2>>','C 2');
CLEANED_NAME := REPLACE(CLEANED_NAME,'D-20=D-20','D-20');
CLEANED_NAME := REPLACE(CLEANED_NAME,'D-4=D-4','D-4');
CLEANED_NAME := REPLACE(CLEANED_NAME,'F-17=F-17','F-17');
CLEANED_NAME := REPLACE(CLEANED_NAME,'E-9=E-9','E-9');
CLEANED_NAME := REPLACE(CLEANED_NAME,'XXII=12','XXII');
CLEANED_NAME := REPLACE(CLEANED_NAME,'XXIII=23','XXIII');
CLEANED_NAME := REPLACE(CLEANED_NAME,'XXIV=24','XXIV');
CLEANED_NAME := REPLACE(CLEANED_NAME,'V=5','V');
CLEANED_NAME := REPLACE(CLEANED_NAME,'6th=6','6th');
CLEANED_NAME := REPLACE(CLEANED_NAME,'2nd=2','2nd');
CLEANED_NAME := REPLACE(CLEANED_NAME,'3as=3','3as');
CLEANED_NAME := REPLACE(CLEANED_NAME,'7=7','7');
CLEANED_NAME := REPLACE(CLEANED_NAME,'8=8','8');
CLEANED_NAME := REPLACE(CLEANED_NAME,'T331=T331','T331');
CLEANED_NAME := REPLACE(CLEANED_NAME,'D20.23=D20.23','D20.23');
CLEANED_NAME := REPLACE(CLEANED_NAME,'<');
CLEANED_NAME := REPLACE(CLEANED_NAME,'>');
CLEANED_NAME := RTRIM(CLEANED_NAME,',');
CLEANED_NAME := TRIM(CLEANED_NAME);
UPDATE AUTOR_TAB SET AUTOR_NORMALIZADO=CLEANED_NAME WHERE AUTOR_TAB.ID =
REC_CURRENT_AUTOR.ID;
COMMIT;
EXCEPTION
WHEN NO_DATA_FOUND
THEN
    DBMS_OUTPUT.PUT_LINE('[[DEBUG]Autor '||REC_CURRENT_AUTOR.ID||' não encontrado');
END;
END LOOP;
EXCEPTION
WHEN OTHERS THEN
    DBMS_OUTPUT.PUT_LINE('Operation failed SQLCODE: '||SQLCODE);
    DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
END;
-- End of NORMALIZA_AUTORES;
END "NORMALIZA_AUTORES";

```

I.5 Código para associar dimensão “Documento” à dimensão “Classe”

```
CREATE OR REPLACE
PACKAGE BODY "FILL_CLASSE_ATTR" AS

PROCEDURE "FILL_CLASSE_ATTR_CDU" IS
--initialize variables here
CURSOR CLASSES_CURSOR IS
  SELECT * FROM CLASSE_TAB WHERE CN3 IS NOT NULL AND CN0 = 'CDU' ORDER BY LENGTH(CN3), CN3 ASC;

REC_CURRENT_CLASS CLASSE_TAB%ROWTYPE;
V_NO_CLASSE_DK    CLASSE_TAB.DIMENSION_KEY%TYPE;

-- main body
BEGIN
  DBMS_OUTPUT.ENABLE(NULL);

  OPEN CLASSES_CURSOR;

  LOOP
    FETCH CLASSES_CURSOR INTO REC_CURRENT_CLASS;
    EXIT WHEN CLASSES_CURSOR%NOTFOUND;

    BEGIN
      DBMS_OUTPUT.PUT_LINE(['[DEBUG] A atualizar documentos para a Classe: '||REC_CURRENT_CLASS.CN3];
      UPDATE DOCUMENTO_TAB
      SET CLASSE = REC_CURRENT_CLASS.DIMENSION_KEY
      WHERE CLASSIFICACAO LIKE REC_CURRENT_CLASS.CN3||'%';
      COMMIT;
    EXCEPTION
      WHEN NO_DATA_FOUND
      THEN
        DBMS_OUTPUT.PUT_LINE(['[DEBUG] não foram atualizados']);
    END;
  END LOOP;
  DBMS_OUTPUT.PUT_LINE(['[DEBUG] A atualizar restantes documentos para Classe por omissão...']);

  SELECT DIMENSION_KEY INTO V_NO_CLASSE_DK FROM CLASSE_TAB WHERE CN0='CDU' and CN3 = 'xxxx';

  UPDATE DOCUMENTO_TAB
  SET CLASSE = V_NO_CLASSE_DK
  WHERE
  CLASSE IS NULL AND
  DOCUMENTO_NATUREZA = 'bib'
  ;
  COMMIT;

  CLOSE CLASSES_CURSOR;

  EXCEPTION
    WHEN OTHERS THEN
      DBMS_OUTPUT.PUT_LINE('Operation failed '||SQLCODE: '||SQLCODE);
      DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
-- enter any exception code here
END;
-- End of FILL_CLASSE_ATTR_CDU;
```

```

PROCEDURE "FILL_CLASSE_ATTR_ISAD" IS
--initialize variables here
CURSOR CLASSES_CURSOR IS
  SELECT * FROM CLASSE_TAB WHERE CN3 IS NOT NULL AND CN0 = 'ISAD' ORDER BY LENGTH(CN3), CN3 ASC;

REC_CURRENT_CLASS CLASSE_TAB%ROWTYPE;
V_NO_CLASSE_DK CLASSE_TAB.DIMENSION_KEY%TYPE;

-- main body
BEGIN
  DBMS_OUTPUT.ENABLE(NULL);

  OPEN CLASSES_CURSOR;

  LOOP
    FETCH CLASSES_CURSOR INTO REC_CURRENT_CLASS;
    EXIT WHEN CLASSES_CURSOR%NOTFOUND;

    BEGIN
      DBMS_OUTPUT.PUT_LINE('[[DEBUG] A atualizar documentos para a Classe: '||REC_CURRENT_CLASS.CN3);
      UPDATE DOCUMENTO_TAB
      SET CLASSE = REC_CURRENT_CLASS.DIMENSION_KEY
      WHERE CLASSIFICACAO like "||REC_CURRENT_CLASS.CN3||"%"||";
      COMMIT;
    EXCEPTION
      WHEN NO_DATA_FOUND
      THEN
        DBMS_OUTPUT.PUT_LINE('[[DEBUG] não foram atualizados');
    END;
  END LOOP;
  DBMS_OUTPUT.PUT_LINE('[[DEBUG] A atualizar restantes documentos para Classe por omissão...');

  SELECT DIMENSION_KEY INTO V_NO_CLASSE_DK FROM CLASSE_TAB WHERE CN0='ISAD' and CN3 = 'xxx';

  UPDATE DOCUMENTO_TAB
  SET CLASSE = V_NO_CLASSE_DK
  WHERE
  CLASSE IS NULL AND
  DOCUMENTO_NATUREZA = 'arquivo'
  ;
  COMMIT;

  CLOSE CLASSES_CURSOR;

  EXCEPTION
    WHEN OTHERS THEN
      DBMS_OUTPUT.PUT_LINE('Operation failed '||"SQLCODE: "||SQLCODE);
      DBMS_OUTPUT.PUT_LINE('SQL Error Message '||SQLERRM);
-- enter any exception code here
END;
-- End of FILL_CLASSE_ATTR_ISAD;

END "FILL_CLASSE_ATTR";

```

Índice Remissivo

A

ACID, *xxi*, 26, 91

ARC, *xxi*, 15, 16, 17, 29, 30

D

data mart, 34

DML, *xxi*

DREW, *xxi*, 9, 10

B

bibliomining, 6, 9

E

ERP, *xxi*, 24

ETL, *xxi*, 16, 17, 24, 26, 28, 32

C

Cognos, 16, 17, 29, 30

COUNTER, *xxi*, 8, 19, 20, 21, 31, 96, 97

G

GIS, *xxi*, 27

M

metodologia, 2, 7, 13, 33, 34, 42, 83
Microsoft, 13, 22, 23, 24, 25
MySQL, 22, 27, 28, 29, 91, 92

I

ISO 11620, 13, 15, 30, 95
ISSN, *xxi*, 20

N

Norma, 13, 30, 95, 96
noSQL, *xxi*, 24

J

JUR, *xxi*, 19, 20, 22

O

OLAP, *xxi*, 9, 12, 23, 25
OLEDB, *xxi*, 13
OPAC, *xxi*, 16
Oracle, 13, 23, 24, 27, 28, 91

K

Kettle, 26, 28, 29
KPI, *xxi*, 25

P

PostgreSQL, 22, 26, 28, 29

L

LDAP, *xxi*, 27

R

RAD, *xxi*, 13
RDBMS, *xxi*, 24

SGBD, *xxi*, 13
SQL, *xxi*, 13, 16, 23, 24, 25, 26, 28, 29, 91, 92
SSD, *xxi*, 1, 5, 22
SUSHI, *xxi*, 21

S

SaaS, *xxi*, 18
SCD, *xxi*
SDI, *xxi*, 1, 2, 3, 32, 35, 36, 37, 38, 40, 41, 42,
83, 106, 107, 112

U

UDM, *xxi*, 25
UML, *xxi*, 42
USAS, 20, 22